REMARKS

Claims 128-145 are cancelled herein and it is respectfully requested that claims 146-161 be entered. The first page of the present application incorporates by reference U.S. Patent Application Serial No. 08/027,146 (the '146 application), which was filed 5 March 1993. A substitute specification is submitted herewith as Exhibit A which incorporates the text from the '146 application. A marked version of the substitute specification is attached herewith as Exhibit B. No new matter has been added. In addition, a copy of the '146 application as filed is also attached for reference. Citations to the text below refer to the '146 application.

The new claims are fully described in the application. The '146 application as filed describes a process in which potential epitopes are identified by scanning the amino acid sequence of an antigen using a motif and then synthesizing the epitope sequence on page 10, lines 23-26. Also, nucleotide sequences which encode peptides comprising an HLA-A2.1 epitope are described on page 19, line 17 to page 20, line 13. The specification also describes methods for testing whether peptide fragments bind an HLA-A2.1 molecule and induce a CTL on page 36, line 1 to page 38, line 24 and on page 72, line 35 to page 76, line 30. Peptide epitopes consisting of 9 to 10 amino acids are described throughout the specification, for example, on page 4, lines 31-36. Peptide fragments consisting of 9 to 10 amino acids are described on page 3, line 4 and line 33, and peptides having a length of less than 15 amino acids are described on page 3, lines 31-32. Amino acid anchors at position 2 and the carboxyl terminus of the epitope are set forth in Table 5 on page 42; L, M, V, I, T and A are all tolerated at position 2 (see also page 39, line 29) and V, L, I, A and M are tolerated at the C-terminus (see also, page 39, line 34). The application also discloses that certain amino acids are preferably not permitted in positions 1, 3, 6 and 7 in instances where the binding motif is a 9-mer and certain amino acids are preferred in positions 1, 3, 4, 5 and 7 of these 9-mers. These data are shown in Table 8 on page 48. They form the basis for claims 147-148 and 156-157. Similarly, amino acids which are preferably not permitted in positions 1, 3-5, and 7-9 of 10-mers, and amino acids that are preferred in positions 1, 3, 4, 6 and 8 of 10-mers are set forth in Table 13 on page 55.

This table forms the basis for claims 86-87 and 96-97. Support for claims 152-153 is found on page 10, line 32-page 11, line 9. Support for the particular peptide in claim 161 is found in the present application (not the '146 application) in Table 3 on page 41. Claim 160 is consistent with the support for tolerated amino acids at anchor positions 2 and the C-terminus in 9-mers and 10-mers but does not include any coincidentally disclosed peptides of which applicants are aware.

Thus, the claims do not add new matter as they find a full description in present application, including the '146 specification filed 5 March 1993.

Applicants appreciate that considerable time and effort has been expended in arriving at the appropriate subject matter to be examined. It is believed that the proposed claims are consistent with the restriction requirement. Election between claims drawn to peptides and claims drawn to nucleic acids was required. The claims as presently drawn relate to peptides.

The requirement for a species election is requested to be withdrawn in view of the present submission. The claims do not relate to a different invention, and it is believed that prosecution is advanced by focussing on the appropriate aspects of the invention. Linkage to a T helper epitope is no longer included in the claims; any particular peptide species misses the point of the invention. The consideration of the Examiner in reviewing the explanation of the invention set forth below and in examining the claims as presented is respectfully requested.

The Invention

The invention provides at least two advances in the art of designing immunogenic peptides and the nucleic acids encoding them. The application is based on an article published by the inventors subsequent to the filing of the '146 application, Ruppert, J., *et al., Cell* (1993) 74:929-937. A copy is enclosed as Exhibit C. First, the invention expands the possible candidates by identifying additional residues that are tolerated in the primary anchor positions at position 2 and the C-terminus of the immunogenic peptide. Previous studies by Falk, *et al., Nature* (1991) 351:290-296 and by Hunt, *et al., Science* (1992) 255:1261-1266 had identified an

HLA-A2.1 motif defined as L and M in position 2 and L, V or I in position 9. The inventors have demonstrated, as shown in the specification, that in addition, immunogenic peptides may include I, V, A or T in the 2 position and may include A and M at the C-terminus. Claims 146 and 154 are directed to taking advantage of these newly discovered motifs; it will be noted that in these claims, either A or M must be at the C-terminus or I, V, A or T must be at the 2 position. The specific peptides of claim 160 have been verified to be immunogenic.

In addition, the invention has contributed an understanding of a way to enhance the predictability of the immunogenicity of the peptides by recognizing the importance of, and identifying the nature of, the secondary anchors. This aspect is claimed in dependent claims 147-151 and 156-158.

Thus, additional immunogenic peptides which would not have been contemplated from the knowledge of the art have been found. This is verified in the declaration of Dr. Sette attached hereto. In addition, the amount of routine screening needed to verify immunogenicity has been diminished by the identification of the secondary residues.

Claim 160 is directed to peptides of less than 15 amino acids which are described by the newly discovered primary anchor generated motifs; claim 161 claims a specific peptide.

The following comments address the specific rejections and objections raised by the Office.

Rejection of Claims Under 35 U.S.C. §112, First Paragraph, for Alleged Lack of Written Description Support in the Specification

A number of claims were rejected under 35 U.S.C. §112, first paragraph as the specification allegedly did not provide an adequate written description for those claims. It is submitted that the rejection is not applicable to the presently submitted claims. Location in the application of this support is described above. Furthermore, the new claims lack the terms objected to in the Office action. For example, the new claims do not include the term "epitope

consisting of about 8-11 amino acids" and "Pan DR" epitope. Thus, it is submitted that the new claims find a written description in the application.

Rejection of Claims Under 35 U.S.C. §112, First Paragraph, for Alleged Lack of Possession of the Invention

Claims 128,129,137, and 145 were rejected under 35 U.S.C. §112, first paragraph, as the Office asserts that the subject matter was not described so as reasonably to convey to own skilled in the art that the inventors had possession of the invention - *i.e.*, that although the claimed peptides were described as "immunogenic" in reality there is no assurance that they are. The Office asserts that it would take "undue experimentation" to determine which peptides encompassed by the formulas in the claims actually are immunogenic. It is this point that applicants dispute.

The Office bases its view on several cited documents. Celis is cited as teaching that an immunogenicity assay *per se* is needed to establish peptide immunogenicity because besides MHC binding, other factors such as antigen processing, peptide transport and the composition of the T cell repertoire could determine whether the peptides can be effective as CTL antigens. Rammansee is quoted as teaching the inadequacy of MHC peptide binding assays and Ochoa-Garay, similarly is said to state that variables such as CTL precursor frequency, peptide hydrophobicity and stability can influence even the *in vitro* induction of CTL responses.

There is no question that not all peptides which bind to an MHC Class I antigen will elicit a CTL response restricted by that Class I antigen and that additional assays would be required to verify immunogenicity. The point is that such verification is routine and was routine at the time the application from which priority was claimed was filed. The disclosures of Celis, Ochoa-Garay, and Rammansee do not contradict this. Indeed, Celis is supportive of the points made here by stating on page 1424:

> Several observations support the notion that the immunogenic potential of a peptide is directly correlated with the peptide's binding affinity to the MHC molecule (citations omitted).

Celis further states:

> The HLA-A motifs mentioned above have been recently validated by showing that using these motifs, one can predict the majority of the peptides that exhibit a significant binding potential to its respective MHC molecule. (Citation omitted.)

Thus, the Celis document itself supports the nexus between immunogenic character and HLA binding and the nexus between a particular motif and HLA binding.

The Sette Declaration, provided herewith, demonstrates that as of March 1993, only routine experimentation would be required to identify immunogenic peptides having an HLA-A2.1 motif since methods for screening any particular peptide for CTL induction and recognition were well known and routinely practiced. In addition, it was understood that it would not be expected that every single peptide tested would prove to be immunogenic and this is considered acceptable. This is also consistent with the legal standard enunciated by *In re Wands*, discussed below, which specifically takes account of the "Foreman factors" alluded to by the Office in the citation of *Ex parte Foreman*. Thus, the specification enables the methods of claims 146-153 and 154-159 as well as the peptides of claims 160-161 without requiring undue experimentation.

In addition, it has been established that 9-mers and 10-mers lacking the expanded motif described herein have zero possibility of binding HLA-A2.1 molecules. This is shown, for example, in figure 1 of the Ruppert, *et al.*, article attached hereto. Thus, by confining further experimentation to peptides comprising the expanded motif, considerable screening, albeit routine screening, can be avoided.

As further discussed below, the Board of Patent Appeals and the Court of Appeals for the Federal Circuit have articulated a standard for enablement which does not require a high degree of predictability when routine screening methods are available. Thus, it is clear that the teachings of the specification and the teachings in the art were sufficient for allowing the skilled artisan to routinely practice the claimed subject matter as of March 1993.

A. Declaration of Alessandro Sette

Applicants request that the Examiner refer to the Declaration by Alessandro Sette for the following discussion. Attached to Dr. Sette's declaration is a table showing later published results employing the expanded motif. The documents set forth in the table are submitted along with a PTO 1449 form listing them. Also listed on the PTO 1449 form and submitted with it are the papers listed in paragraph 2 of Dr. Sette's declaration.

The Sette Declaration explains that numerous publications predating the March 1993 effective filing date describe the routine use of both HLA binding assays for identifying potential immunogens (paragraph 6), as well as *in vitro* and *in vivo* methods for confirming whether a peptide is immunogenic or not immunogenic (paragraphs 4 and 5). Exemplary studies in which researchers routinely utilized these assays are described in the declaration in paragraphs 7 and 8.

The declaration also demonstrates that the motif itself reduces the amount of experimentation required to determine whether a peptide is immunogenic. Knowledge of a particular motif, such as the HLA-A2.1 motifs described in the specification, can reduce the number of peptides required for testing in immunogenicity assays by 10.8-fold (paragraph 9). Furthermore, the declaration makes it clear that the skilled artisan was prepared as of March 1993 to screen multiple peptides to determine which of those were immunogenic (paragraph 12, referring to Hill), and that the procedures utilized by the skilled artisan before the filing date of the application are similar to those described in the application (paragraph 13). The approach of using the tolerated motifs to screen for and confirm immunogenic peptides was repeatedly confirmed subsequent to March of 1993 as shown in paragraphs 15-16.

Moreover, because the peptides described in paragraph 11 of the declaration were recognized by CTLs from infected patients, immunized individuals, and naturally exposed individuals, the factors noted by the Office as being potentially problematic are not major obstacles to practicing the claimed subject matter (paragraph 17).

Thus, the Sette declaration demonstrates that the claimed methods can be practiced without undue experimentation due to the predictive value of the HLA-A2.1 motif and the routine application of screening procedures taught in the '146 specification.

B. Post-Filing Evidence that the Claimed Subject Matter is Enabled

Paragraphs 15-16 of the Sette declaration describe studies performed after March 1993 in which researchers utilized the HLA-A2.1 motifs and methods disclosed in the specification to identify immunogenic peptides. As can be seen in the table attached to the declaration, many researchers utilized these motifs to identify multiple immunogenic epitopes from a wide variety of antigens. It is therefore clear that the methods set forth in the specification and claimed herein could be routinely practiced by the skilled artisan without undue experimentation at the time the application was filed.

C. Peptide Processing and Peptide Length Are Not Significant Obstacles to Practicing the Claimed Subject Matter

As noted above, paragraph 17 in the Sette declaration establishes that antigen processing and other factors mentioned in the Office action are not significant obstacles to practicing the claimed subject matter. In addition, other studies provide further evidence that these factors are not a significant barriers to practicing the claimed subject matter.

CTL responses are induced as a consequence of naïve T cells recognizing a complex between antigenic peptides and class I molecules. These peptides, which range in length from 8-11 amino acids, are the result of proteolytic degradation of intact antigen (Niedermann, *et al.*, *Immunol. Rev.* (1999) 172:29-48). These processed peptides then bind to class I molecules and are presented on the cell surface of the cell. While longer fragments of the antigen may be utilized to induce CTL responses, ultimately it is the presentation of the minimal epitope by the corresponding class I molecule that results in the differentiation and expansion of the naïve T cell.

Further, the binding of the minimal epitope is dictated by the sequences within the epitope itself and not the surrounding amino acids (Sette & Sidney, *Curr. Opin. Immunol.* (1998) 10:478-482; Medden, *Ann. Rev. Immunol.* (1995) 13:587), and therefore, the motif set forth in the claimed methods provides for HLA binding and the resulting immune response. In the case of longer peptides, the presence of additional amino acids does prevent complex formation, but these epitopes are typically processed by intracellular proteosomes that recognize cleavage sites adjacent to the minimal epitope (Del Val, *et al.*, *Cell* (1991) 66:1145-1153; Eisenlohr, *et al.*, *J. Exp. Med.* (1992) 175:481-487). Using variable length fragments, it has been demonstrated that efficient epitope processing can occur irrespective of the fragment's length (Niedermann, *et al.*, *Proc. Natl. Acad. Sci. USA* (1996) 93:8572-8577; Niedermann, *et al.*, *Immunol. Rev.* (1999) 172:29-48).

Additionally, epitope processing is not limited exclusively to intracellular proteosomes. Several studies have shown that serum and membrane associated proteases can effectively process longer peptides resulting in the generation of the minimal epitope capable of being presented bound to the corresponding HLA class I molecule (Sherman, *et al.*, *J. Exp. Med.* (1992) 175:1221-1226; Kozlowski, *et al.*, *J. Exp. Med.* (1992) 175:1417-1422). Thus, the presence of additional residues besides the minimal epitope might influence the degree of immunogenicity or antigenicity, defined as the amount of peptide necessary to achieve a given level of response. However, antigen length is not a major variable with regard to practicing the claimed methods in the sense that longer peptides are clearly active in various *in vitro* and *in vivo* models (see for example Kast, *et al.*, *Eur. J. Immunol.* (1993) 23:1189-1192).

Moreover, other studies have demonstrated that multi-epitope polypeptides and minigenes can be effectively used to induce cellular immune responses, in particular by enhancing the efficiency of epitope processing. Binding of the optimal epitopes remains constant. In fact, armed with the knowledge of the certain CTL-inducing epitopes, researchers were able to augment immune responses. For example, Shastri, *et al.*, *J. Immunol.* (1995) 155:4339-4346 and Bergmann, *et al.*, *J. Virol.* (1994) 68:5306-5310 demonstrated increasingly

efficient epitope processing resulting from changing the amino acids that flank an epitope. Moreover, Ishioka, *et al.*, *J. Immunol.* (1999) 162:3915-3925 demonstrated that it is possible to induce CTL responses ***that are stronger*** than those induced by an intact antigen when multiple epitopes are delivered as a string of peptides.

These examples further illustrate that antigen processing and other factors do not significantly impact binding between the epitope and the HLA class I molecule or impact the immunogenicity of the component epitopes once generated. Thus, applicants submit that claims 146-150, 154, and 156-161, which for example may require the peptide to be processed from a longer polypeptide before contacting the HLA molecules, are fully enabled.

D. The Rejection is Analogous to the Rejection at Issue in *In re Wands*

The routine screening procedure for identifying CTL-inducing peptides is analogous to a screening procedure that the Federal Circuit deemed did not require undue experimentation. The claims at issue in *In re Wands*, 8 USPQ.2d 1400 (Fed. Cir. 1988) were directed to an assay method that required IgM antibodies having affinity for a specific antigen. At issue in *Wands* was whether screening a large number of antibodies to select those required represented undue experimentation. Because the present rejection raises this precise issue - *i.e.*, whether screening multiple peptides to select those that are immunogenic represents undue experimentation, the holding in *Wands* is believed controlling.

The Court in *Wands* determined that screening multiple hybridomas to select particular monoclonal antibodies did not require undue experimentation because 1) screening procedures were taught in the specification and 2) researchers in the field were prepared to screen a large number of hybridomas. Thus, while these hybridoma screening procedures required a significant amount of experimentation, the procedures were routine and the claims were therefore in accordance with 35 U.S.C. §112, first paragraph.

The *Wands* Court noted that Wands used a commercially available radioimmunoassay kit to screen clones in a preliminary screen which identified a subset of candidates that, in order to

satisfy the limitation of the claims, "require further screening to select those which have an IgM isotype and have a binding affinity constant of at least 10^9 M." The Court noted that the results of the test kit do not provide a numerical affinity constant, which must be measured using the more laborious Schatchard analysis. Similarly here, the specification teaches several stages of testing whereby a particular motif is first selected, binding to the appropriate HLA-A2.1 antigen is used to identify a subset that binds and this subset is used in the slightly more laborious assay for immunogenicity. As demonstrated in the Sette declaration, a person of ordinary skill in the art is prepared to screen a number of peptides that correspond to a motif.

Thus, experimentation for practicing the claimed methods and making the claimed nucleic acids is not undue because (1) the specification teaches motifs and screening procedures for identifying CTL-inducing peptides, (2) the screening procedures can be practiced in a routine manner, and (3) researchers are prepared to use these methods to screen a large group of peptides. The facts here are directly analogous to those in *Wands*.

E. The Rejection is Analogous to the Rejection at Issue in *Ex parte Mark*

The situation in the present application is also analogous to that in *Ex parte Mark*, 12 USPQ.2d 1905 (BPAI 1989). In particular, the reasoning set forth in the present Office action is similar to the Examiner's reasoning in *Mark*. The relevant claims at issue were directed to a method for producing DNA encoding a synthetic mutein of any protein by substituting other amino acids for cysteine, wherein the mutein had the biological activity of the parent protein. The claims were rejected as being non-enabled, based on the prior art disclosure that eight such muteins of two different proteins lacked or were substantially reduced in the biological activity of the parent proteins. The examiner reasoned that:

> it would require undue further experimentation to construct . . . the *innumerable muteins* encompassed by the instant claims . . . and to screen the muteins produced for any of those which exhibit biological activity after modification.

Id. at 1906 (emphasis added). Additionally, the examiner asserted that most of the muteins "would be inoperative," and that there was an established unpredictability as to how many muteins would have to be produced in order to obtain even one biologically active embodiment.

On appeal, the Board reversed the enablement rejection, holding that the claims were enabled because they all required "that the mutein retain the biological activity of the native protein." *Id.* at 1906-7. The prior art muteins lacking activity were "merely examples of work which is *outside* the claims." *Id.* at 1907. The Board was persuaded that only routine experimentation was required to determine whether a cysteine substitution or replacement would result in a mutein within the scope of the claims. Importantly, the Board stated:

> [t]o the extent that the examiner is concerned that undue experimentation would be required to determine other proteins suitable for use in the present invention, we find . . . that only routine experimentation would be needed for one skilled in the art to practice the claimed invention for a given protein. *The fact that a given protein may not be amenable for use in the present invention* in that the cysteine residues are needed for the biological activity of the protein *does not militate against a conclusion of enablement.* One skilled in the art is clearly enabled to perform such work as needed to determine whether the cysteine residues of a given protein are needed for retention of biological activity.

Id. at 1907 (emphasis added).

The Board found persuasive the fact that (a) muteins having activity could be routinely identified through the methods disclosed in the specification and the general knowledge in the art, and (b) the claimed method had successfully identified three proteins for which muteins could be made that had the required activity.

Mark is entirely analogous. Claim 146 requires that the peptide is immunogenic in a subject comprising an HLA-A2.1 molecule and that it binds an HLA-A2.1 molecule; claim 154 has a similar limitation as does claim 160. Thus, as in *Mark*, any inoperative embodiments are "merely examples of work which is outside the claims." Just as in *Mark*, where "one skilled in the art is clearly enabled to perform such work as needed to determine whether the cysteine residues of a given protein are needed for retention of biological activity," one skilled in the art is

clearly enabled to "perform such work as needed to determine whether" the peptides bearing the required motifs are immunogenic. And just as in *Mark* (and in *Wands*), routine assays are described in the specification. And as in *Mark*, the claimed methods also have successfully identified immunogenic peptides. This is confirmed in the Sette declaration paragraphs 15-16.

F. Conclusion

The Sette Declaration and legal precedent make it clear that the claimed methods and compositions may be practiced without undue experimentation, and the rejection under 35 U.S.C. § 112, first paragraph, may properly be withdrawn.

Rejection Under 35 U.S.C. §112, Second Paragraph

A number of claims were rejected under 35 U.S.C. §112, second paragraph, as the terms "at the carboxyl terminus" and "c-terminal position" in the context of the previous claims were allegedly indefinite. As presented in the new claims, the term "C-terminus" is definite as it is referenced to the subsequence or motif. Accordingly, the rejection under 35 U.S.C. § 112, second paragraph, may be withdrawn.

Rejection Under 35 U.S.C. § 103(a)

Claims 128, 129, and 137 were rejected under 35 U.S.C. §103(a) as allegedly being unpatentable over Boon *et al.*, Cheever *et al.*, or Kubo, *et al.* in view of Sette, *et al.* The new claims are supported in the '146 application which was filed 5 March 1993, so the Sette patent filed 14 September 1993 is not prior art. Thus, on a purely formal basis, the rejection as framed is inapplicable to the new claims.

Furthermore, a *prima facie* case for obvious may not be maintained because the cited documents do not result in the claimed subject matter. Claims 146-161 require the steps of (a) providing an amino acid sequence of an antigen of interest, (b) identifying a specific subsequence within that antigen characterized by the motifs described in step (b), and (c) identifying a fragment of the antigen which contains this subsequence. Neither Boon, Cheever,

nor Kubo carries out this set of steps. In particular, neither document discloses identifying subsequence as required by step (b).

Boon describes cloning experiments leading to the nucleic acid sequence of MAGE-1 (column 5) and particular CTL epitopes known in the art (see for example column 14) which is an entirely different method. There is no consideration of identifying a motif specific for a particular HLA antigen, much less HLA-A2.1. Thus, step (b) of claim 146 is completely missing and therefore is not taught or suggested by Boon.

Similarly, Cheever fails to carry out step (b). Instead, Cheever is concerned with identifying peptides derived from the Her2/Neu oncogene that induce an antibody response. There is no teaching or suggestion of scanning the amino acid sequence of the Her2/Neu oncogene for the motif in step (b).

Furthermore, Kubo does not mention the motif set forth in step (b). While Kubo describes methods for scanning antigens for subsequences which correspond to a particular motif, there is no teaching or suggestion of utilizing the particular motifs set forth in the claims. Thus, a teaching or suggestion of step (b) is also missing from Kubo.

Thus, the cited documents fail to describe a method for identifying a specific subsequence within that antigen characterized by the motifs described in step (b), and also fail to describe the motif required by all of the claims. The cited documents therefore may not establish a basis for a *prima facie* case of obviousness.

Conclusions

Formal rejections have been addressed by amendment, and a rejection for lack of an enabling written description has been shown by the declaration of Dr. Sette to be misplaced. As demonstrated by declaratory evidence, only routine experimentation is required to determine which peptides fall within the scope of the claims. Such routine experimentation is recognized as acceptable by the holdings in *In re Wands* and *Ex parte Mark*. Further, no suggestion of the claims as presently proposed is found in the cited documents. For these reasons, it is believed

that claims 146-161 are in a position for allowance and passage of these claims to issue is respectfully requested.

In the unlikely event that the transmittal letter is separated from this document and the Patent Office determines that an extension and/or other relief is required, applicants petition for any required relief including extensions of time and authorize the Commissioner to charge the cost of such petitions and/or other fees due in connection with the filing of this document to **Deposit Account No. 03-1952** referencing docket No. 399632000623.

Respectfully submitted,

Dated: November 15, 2001

By: _____

Bruce Grant
Registration No. 47,608

Morrison & Foerster LLP
3811 Valley Centre Drive,
Suite 500
San Diego, California 92130-2332
Telephone: (858) 720-7962
Facsimile: (858) 720-5125

HLA-A2.1 BINDING PEPTIDES AND THEIR USES

Cross-Reference to Related Applications

[0001] The present application is a continuation in part of USSN 08/159,184, filed 29 November 1993, Gray, et al. (abandoned), which is a continuation in part of USSN 08/073,205, filed 4 June 1993, Gray et al. (abandoned), which is a continuation in part of 08/027,146, filed 5 March 1993, Gray, et al. (abandoned). [It is related to USSN 08/205,13.] All of which are incorporated herein by reference.

Technical Field

[0002] The present invention relates to compositions and methods for preventing, treating or diagnosing a number of pathological states such as viral diseases and cancers. In particular, it provides novel peptides capable of binding selected major histocompatibility complex (MHC) molecules and inducing an immune response.

Background Art

[0003] MHC molecules are classified as either Class I or Class II molecules. Class II MHC molecules are expressed primarily on cells involved in initiating and sustaining immune responses, such as T lymphocytes, B lymphocytes, macrophages, etc. Class II MHC molecules are recognized by helper T lymphocytes and induce proliferation of helper T lymphocytes and amplification of the immune response to the particular immunogenic peptide that is displayed. Class I MHC molecules are expressed on almost all nucleated cells and are recognized by cytotoxic T lymphocytes (CTLs), which then destroy the antigen-bearing cells. CTLs are particularly important in tumor rejection and in fighting viral infections. The CTL recognizes the antigen in the form of a peptide fragment bound to the MHC class I molecules rather than the intact foreign antigen itself. The antigen must normally be endogenously synthesized by the cell, and a portion of the protein antigen is degraded into small peptide fragments in the cytoplasm. Some of these small peptides translocate into a pre-Golgi compartment and interact with class I heavy chains to facilitate proper folding and association with the subunit $\alpha 2$ microglobulin. The peptide-MHC class I

Exhibit B

complex is then routed to the cell surface for expression and potential recognition by specific CTLs.

[0004] Investigations of the crystal structure of the human MHC class I molecule, HLA-A2.1, indicate that a peptide binding groove is created by the folding of the $\alpha 1$ and $\alpha 2$ domains of the class I heavy chain (Bjorkman et al., Nature 329:506 (1987). In these investigations, however, the identity of peptides bound to the groove was not determined.

[0005] Buus et al., Science 242:1065 (1988) first described a method for acid elution of bound peptides from MHC. Subsequently, Rammensee and his coworkers (Falk et al., Nature 351:290 (1991) have developed an approach to characterize naturally processed peptides bound to class I molecules. Other investigators have successfully achieved direct amino acid sequencing of the more abundant peptides in various HPLC fractions by conventional automated sequencing of peptides eluted from class I molecules of the B type (Jardetzky, et al., Nature 353:326 (1991) and of the A2.1 type by mass spectrometry (Hunt, et al., Science 225:1261 (1992). A review of the characterization of naturally processed peptides in MHC Class I has been presented by Rötzschke and Falk (Rötzschke and Falk, Immunol. Today 12:447 (1991).

[0006] Sette et al., Proc. Natl. Acad. Sci. USA 86:3296 (1989) showed that MHC allele specific motifs could be used to predict MHC binding capacity. Schaeffer et al., Proc. Natl. Acad. Sci. USA 86:4649 (1989) showed that MHC binding was related to immunogenicity. Several authors (De Bruijn et al., Eur. J. Immunol., 21:2963-2970 (1991); Pamer et al., 991 Nature 353:852-955 (1991)) have provided preliminary evidence that class I binding motifs can be applied to the identification of potential immunogenic peptides in animal models. Class I motifs specific for a number of human alleles of a given class I isotype have yet to be described. It is desirable that the combined frequencies of these different alleles should be high enough to cover a large fraction or perhaps the majority of the human outbred population.

[0007] Despite the developments in the art, the prior art has yet to provide a useful human peptide-based vaccine or therapeutic agent based on this work. The present invention provides these and other advantages.

Disclosure of the Invention

[0008] The present invention provides compositions comprising immunogenic peptides having binding motifs for HLA-A2.1 molecules. The immunogenic peptides, which bind to the appropriate MHC allele, are preferably 9 to 10 residues in length and comprise conserved residues at certain positions such as positions 2 and 9. Moreover, the peptides do not comprise negative binding residues as defined herein at other positions such as positions 1, 3, 6 and/or 7 in the case of peptides 9 amino acids in length and positions 1, 3, 4, 5, 7, 8 and/or 9 in the case of peptides 10 amino acids in length. The present invention defines positions within a motif enabling the selection of peptides which will bind efficiently to HLA A2.1.

[0009] Epitopes on a number of immunogenic target proteins can be identified using the peptides of the invention. Examples of suitable antigens include prostate cancer specific antigen (PSA), hepatitis B core and surface antigens (HBVc, HBVs) hepatitis C antigens, Epstein-Barr virus antigens, human immunodeficiency type-1 virus (HIV1) and papilloma virus antigens. The peptides are thus useful in pharmaceutical compositions for both in vivo and ex vivo therapeutic and diagnostic applications.

Definitions

[0010] The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other typically by peptide bonds between the alpha-amino and carbonyl groups of adjacent amino acids. The oligopeptides of the invention are less than about 15 residues in length and usually consist of between about 8 and about 11 residues, preferably 9 or 10 residues.

[0011] An "immunogenic peptide" is a peptide which comprises an allele-specific motif such that the peptide will bind an MHC molecule and induce a CTL response. Immunogenic peptides of the invention are capable of binding to an appropriate HLA-A2.1 molecule and inducing a cytotoxic T cell response against the antigen from which the immunogenic peptide is derived.

[0012] Immunogenic peptides are conveniently identified using the algorithms of the invention. The algorithms are mathematical procedures that produce a score which enables the selection of immunogenic peptides. Typically one uses the algorithmic score with a "binding threshold" to enable selection of peptides that have a high probability of binding at

a certain affinity and will in turn be immunogenic. The algorithm is based upon either the effects on MHC binding of a particular amino acid at a particular position of a peptide or the effects on binding of a particular substitution in a motif containing peptide.

[0013] A "conserved residue" is an amino acid which occurs in a significantly higher frequency than would be expected by random distribution at a particular position in a peptide. Typically a conserved residue is one where the MHC structure may provide a contact point with the immunogenic peptide. One to three, preferably two, conserved residues within a peptide of defined length defines a motif for an immunogenic peptide. These residues are typically in close contact with the peptide binding groove, with their side chains buried in specific pockets of the groove itself. Typically, an immunogenic peptide will comprise up to three conserved residues, more usually two conserved residues.

[0014] As used herein, "negative binding residues" are amino acids which if present at certain positions (for example, positions 1, 3 and/or 7 of a 9-mer) will result in a peptide being a nonbinder or poor binder and in turn fail to be immunogenic i.e. induce a CTL response.

[0015] The term "motif" refers to the pattern of residues in a peptide of defined length, usually about 8 to about 11 amino acids, which is recognized by a particular MHC allele. The peptide motifs are typically different for each human MHC allele and differ in the pattern of the highly conserved residues and negative residues.

[0016] The binding motif for an allele can be defined with increasing degrees of precision. In one case, all of the conserved residues are present in the correct positions in a peptide and there are no negative residues in positions 1, 3 and/or 7.

[0017] The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany it as found in its native state. Thus, the peptides of this invention do not contain materials normally associated with their in situ environment, e.g., MHC I molecules on antigen presenting cells. Even where a protein has been isolated to a homogenous or dominant band, there are trace contaminants in the range of 5-10% of native protein which co-purify with the desired protein. Isolated peptides of this invention do not contain such endogenous co-purified protein.

[0018] The term "residue" refers to an amino acid or amino acid mimetic incorporated in an oligopeptide by an amide bond or amide bond mimetic.

Description of the Preferred Embodiments

[0019] The present invention relates to the determination of allele-specific peptide motifs for human Class I MHC (sometimes referred to as HLA) allele subtypes, in particular, peptide motifs recognized by HLA-A2.1 alleles. These motifs are then used to define T cell epitopes from any desired antigen, particularly those associated with human viral diseases, cancers or autoiummune diseases, for which the amino acid sequence of the potential antigen or autoantigen targets is known.

[0020] Epitopes on a number of potential target proteins can be identified in this manner. Examples of suitable antigens include prostate specific antigen (PSA), hepatitis B core and surface antigens (HBVc, HBVs) hepatitis C antigens, Epstein-Barr virus antigens, melanoma antigens (e.g., MAGE-1), human immunodeficiency virus (HIV) antigens and human papilloma virus (HPV) antigens.

[0021] Autoimmune associated disorders for which the peptides of the invention may be employed to relieve the symptoms of, treat or prevent the occurrence or reoccurrence of include, for example, multiple sclerosis (MS), rheumatoid arthritis (RA), Sjogren syndrome, scleroderma, polymyositis, dermatomyositis, systemic lupus erythematosus, juvenile rheumatoid arthritis, ankylosing spondylitis, myasthenia gravis (MG), bullous pemphigoid (antibodies to basement membrane at dermal-epidermal junction), pemphigus (antibodies to mucopolysaccharide protein complex or intracellular cement substance), glomerulonephritis (antibodies to glomerular basement membrane), Goodpasture's syndrome, autoimmune hemolytic anemia (antibodies to erythrocytes), Hashimoto's disease (antibodies to thyroid), pernicious anemia (antibodies to intrinsic factor), idiopathic thrombocytopenic purpura (antibodies to platelets), Grave's disease, and Addison's disease (antibodies to thyroglobulin), and the like.

[0022] The autoantigens associated with a number of these diseases have been identified. For example, in experimentally induced autoimmune diseases, antigens involved in pathogenesis have been characterized: in arthritis in rat and mouse, native type-II collagen is identified in collagen-induced arthritis, and mycobacterial heat shock protein in adjuvant arthritis; thyroglobulin has been identified in experimental allergic thyroiditis (EAT) in mouse; acetyl choline receptor (AChR) in experimental allergic myasthenia gravis (EAMG); and myelin basic protein (MBP) and proteolipid protein (PLP) in experimental allergic

encephalomyelitis (EAE) in mouse and rat. In addition, target antigens have been identified in humans: type-II collagen in human rheumatoid arthritis; and acetyl choline receptor in myasthenia gravis.

[0023] Peptides comprising the epitopes from these antigens are synthesized and then tested for their ability to bind to the appropriate MHC molecules in assays using, for example, purified class I molecules and radioiodonated peptides and/or cells expressing empty class I molecules by, for instance, immunofluorescent staining and flow microfluorometry, peptide-dependent class I assembly assays, and inhibition of CTL recognition by peptide competition. Those peptides that bind to the class I molecule are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary in vitro or in vivo CTL responses that can give rise to CTL populations capable of reacting with virally infected target cells or tumor cells as potential therapeutic agents.

[0024] The MHC class I antigens are encoded by the HLA-A, B, and C loci. HLA-A and B antigens are expressed at the cell surface at approximately equal densities, whereas the expression of HLA-C is significantly lower (perhaps as much as 10-fold lower). Each of these loci have a number of alleles. The peptide binding motifs of the invention are relatively specific for each allelic subtype.

[0025] For peptide-based vaccines, the peptides of the present invention preferably comprise a motif recognized by an MHC I molecule having a wide distribution in the human population. Since the MHC alleles occur at different frequencies within different ethnic groups and races, the choice of target MHC allele may depend upon the target population. Table 1 shows the frequency of various alleles at the HLA-A locus products among different races. For instance, the majority of the Caucasoid population can be covered by peptides which bind to four HLA-A allele subtypes, specifically HLA-A2.1, A1, A3.2, and A24.1. Similarly, the majority of the Asian population is encompassed with the addition of peptides binding to a fifth allele HLA-A11.2.

TABLE 1

A Allele/Subtype	N(69)*	A(54)	C(502)
A1	10.1(7)	1.8(1)	27.4(138)
A2.1	11.5(8)	37.0(20)	39.8(199)
A2.2	10.1(7)	0	3.3(17)
A2.3	1.4(1)	5.5(3)	0.8(4)
A2.4	-	-	-
A2.5	-	-	-
A3.1	1.4(1)	0	0.2(0)
A3.2	5.7(4)	5.5(3)	21.5(108)
A11.1	0	5.5(3)	0
A11.2	5.7(4)	31.4(17)	8.7(44)
A11.3	0	3.7(2)	0
A23	4.3(3)	-	3.9(20)
A24	2.9(2)	27.7(15)	15.3(77)
A24.2	-	-	-
A24.3	-	-	-
A25	1.4(1)	-	6.9(35)
A26.1	4.3(3)	9.2(5)	5.9(30)
A26.2	7.2(5)	-	1.0(5)
A26V	-	3.7(2)	-
A28.1	10.1(7)	-	1.6(8)
A28.2	1.4(1)	-	7.5(38)
A29.1	1.4(1)	-	1.4(7)
A29.2	10.1(7)	1.8(1)	5.3(27)
A30.1	8.6(6)	-	4.9(25)
A30.2	1.4(1)	-	0.2(1)
A30.3	7.2(5)	-	3.9(20)
A31	4.3(3)	7.4(4)	6.9(35)
A32	2.8(2)	-	7.1(36)
Aw33.1	8.6(6)	-	2.5(13)
Aw33.2	2.8(2)	16.6(9)	1.2(6)
Aw34.1	1.4(1)	-	-
Aw34.2	14.5(10)	-	0.8(4)
Aw36	5.9(4)	-	-

Table compiled from B. DuPont, Immunobiology of HLA, Vol. I, Histocompatibility Testing 1987, Springer-Verlag, New York 1989.

* N - negroid; A = Asian; C = caucasoid. Numbers in parenthesis represent the number of individuals included in the analysis.

[0026] The nomenclature used to describe peptide compounds follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the carboxyl group to the right (the C-terminus) of each amino acid residue. In the formulae representing

selected specific embodiments of the present invention, the amino- and carboxyl-terminal groups, although not specifically shown, are in the form they would assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or single letter designations. The L-form of an amino acid residue is represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids having D-forms is represented by a lower case single letter or a lower case three letter symbol. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or G.

Modes of Carrying Out the Invention

[0027] The procedures used to identify peptides of the present invention generally follow the methods disclosed in Falk et al., Nature 351:290 (1991), which is incorporated herein by reference. Briefly, the methods involve large-scale isolation of MHC class I molecules, typically by immunoprecipitation or affinity chromatography, from the appropriate cell or cell line. Examples of other methods for isolation of the desired MHC molecule equally well known to the artisan include ion exchange chromatography, lectin chromatography, size exclusion, high performance ligand chromatography, and a combination of all of the above techniques.

[0028] In the typical case, immunoprecipitation is used to isolate the desired allele. A number of protocols can be used, depending upon the specificity of the antibodies used. For example, allele-specific mAb reagents can be used for the affinity purification of the HLA-A, HLA-B_1, and HLA-C molecules. Several mAb reagents for the isolation of HLA-A molecules are available. The monoclonal BB7.2 is suitable for isolating HLA-A2 molecules. Affinity columns prepared with these mAbs using standard techniques are successfully used to purify the respective HLA-A allele products.

[0029] In addition to allele-specific mAbs, broadly reactive anti-HLA-A, B, C mAbs, such as W6/32 and B9.12.1, and one anti-HLA-B, C mAb, B1.23.2, could be used in alternative affinity purification protocols as described in the example section below. The peptides bound to the peptide binding groove of the isolated MHC molecules are eluted typically using acid treatment. Peptides can also be dissociated from class I molecules by a variety of standard denaturing means, such as heat, pH, detergents, salts, chaotropic agents, or a combination thereof. Peptide fractions are further separated from the MHC molecules by

reversed-phase high performance liquid chromatography (HPLC) and sequenced. Peptides can be separated by a variety of other standard means well known to the artisan, including filtration, ultrafiltration, electrophoresis, size chromatography, precipitation with specific antibodies, ion exchange chromatography, isoelectrofocusing, and the like.

[0030] Sequencing of the isolated peptides can be performed according to standard techniques such as Edman degradation (Hunkapiller, M.W., et al., Methods Enzymol. 91, 399 [1983]). Other methods suitable for sequencing include mass spectrometry sequencing of individual peptides as previously described (Hunt, et al., Science 225:1261 (1992), which is incorporated herein by reference). Amino acid sequencing of bulk heterogenous peptides (e.g., pooled HPLC fractions) from different class I molecules typically reveals a characteristic sequence motif for each class I allele.

[0031] Definition of motifs specific for different class I alleles allows the identification of potential peptide epitopes from an antigenic protein whose amino acid sequence is known. Typically, identification of potential peptide epitopes is initially carried out using a computer to scan the amino acid sequence of a desired antigen for the presence of motifs. The epitopic sequences are then synthesized. The capacity to bind MHC Class molecules is measured in a variety of different ways. One means is a Class I molecule binding assay as described in Example 8, below. Other alternatives described in the literature include inhibition of antigen presentation (Sette, et al., J. Immunol. 141:3893 (1991), in vitro assembly assays (Townsend, et al., Cell 62:285 (1990), and FACS based assays using mutated ells, such as RMA.S (Melief, et al., Eur. J. Immunol. 21:2963 (1991)).

[0032] Next, peptides that test positive in the MHC class I binding assay are assayed for the ability of the peptides to induce specific CTL responses in vitro. For instance, [A]antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells (Inaba, et al., J. Exp. Med. 166:182 (1987); Boog, Eur. J. Immunol. 18:219 [1988]).

[0033] Alternatively, mutant mammalian cell lines that are deficient in their ability to load class I molecules with internally processed peptides, such as the mouse cell lines RMA-S (Kärre, et al.. Nature, 319:675 (1986); Ljunggren, et al., Eur. J. Immunol. 21:2963-2970 (1991)), and the human somatic T cell hybrid, T-2 (Cerundolo, et al., Nature 345:449-452 (1990)) and which have been transfected with the appropriate human class I genes are

conveniently used, when peptide is added to them, to test for the capacity of the peptide to induce in vitro primary CTL responses. Other eukaryotic cell lines which could be used include various insect cell lines such as mosquito larvae (ATCC cell lines CCL 125, 126, 1660, 1591, 6585, 6586), silkworm (ATTC CRL 8851), armyworm (ATCC CRL 1711), moth (ATCC CCL 80) and Drosophila cell lines such as a Schneider cell line (see Schneider J. Embryol. Exp. Morphol. 27:353-365 [1927]).

[0034] Peripheral blood lymphocytes are conveniently isolated following simple venipuncture or leukapheresis of normal donors or patients and used as the responder cell sources of CTL precursors. In one embodiment, the appropriate antigen-presenting cells are incubated with 10-100 μM of peptide in serum-free media for 4 hours under appropriate culture conditions. The peptide-loaded antigen-presenting cells are then incubated with their responder cell populations in vitro for 7 to 10 days under optimized culture conditions. Positive CTL activation can be determined by assaying the cultures for the presence of CTLs that kill radiolabeled target cells, both specific peptide-pulsed targets as well as target cells expressing endogenously processed form of the relevant virus or tumor antigen from which the peptide sequence was derived.

[0035] Specificity and MHC restriction of the CTL is determined by testing against different peptide target cells expressing appropriate or inappropriate human MHC class I. The peptides that test positive in the MHC binding assays and give rise to specific CTL responses are referred to herein as immunogenic peptides.

[0036] The immunogenic peptides can be prepared synthetically, or by recombinant DNA technology or from natural sources such as whole viruses or tumors. Although the peptide will preferably be substantially free of other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides can be synthetically conjugated to native fragments or particles.

[0037] The polypeptides or peptides can be a variety of lengths, either in their neutral (uncharged) forms or in forms which are salts, and either free of modifications such as glycosylation, side chain oxidation, or phosphorylation or containing these modifications, subject to the condition that the modification not destroy the biological activity of the polypeptides as herein described.

[0038] Desirably, the peptide will be as small as possible while still maintaining substantially all of the biological activity of the large peptide. When possible, it may be

desirable to optimize peptides of the invention to a length of 9 or 10 amino acid residues, commensurate in size with endogenously processed viral peptides or tumor cell peptides that are bound to MHC class I molecules on the cell surface.

[0039] Peptides having the desired activity may be modified as necessary to provide certain desired attributes, e.g., improved pharmacological characteristics, while increasing or at least retaining substantially all of the biological activity of the unmodified peptide to bind the desired MHC molecule and activate the appropriate T cell. For instance, the peptides may be subject to various changes, such as substitutions, either conservative or non-conservative, where such changes might provide for certain advantages in their use, such as improved MHC binding. By conservative substitutions is meant replacing an amino acid residue with another which is biologically and/or chemically similar, e.g., one hydrophobic residue for another, or one polar residue for another. The substitutions include combinations such as Gly, Ala; Val, Ile, Leu, Met; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. The effect of single amino acid substitutions may also be probed using D-amino acids. Such modifications may be made using well known peptide synthesis procedures, as described in e.g., Merrifield, Science 232:341-347 (1986), Barany and Merrifield, The Peptides, Gross and Meienhofer, eds. (N.Y., Academic Press), pp. 1-284 (1979); and Stewart and Young, Solid Phase Peptide Synthesis, (Rockford, Ill., Pierce), 2d Ed. (1984), incorporated by reference herein.

[0040] The peptides can also be modified by extending or decreasing the compound's amino acid sequence, e.g., by the addition or deletion of amino acids. The peptides or analogs of the invention can also be modified by altering the order or composition of certain residues, it being readily appreciated that certain amino acid residues essential for biological activity, e.g., those at critical contact sites or conserved residues, may generally not be altered without an adverse effect on biological activity. The non-critical amino acids need not be limited to those naturally occurring in proteins, such as L-α-amino acids, or their D-isomers, but may include non-natural amino acids as well, such as ß-δ-8-amino acids, as well as many derivatives of L-α-amino acids.

[0041] Typically, a series of peptides with single amino acid substitutions are employed to determine the effect of electrostatic charge, hydrophobicity, etc. on binding. For instance, a series of positively charged (e.g., Lys or Arg) or negatively charged (e.g., Glu) amino acid substitutions are made along the length of the peptide revealing different patterns of

sensitivity towards various MHC molecules and T cell receptors. In addition, multiple substitutions using small, relatively neutral moieties such as Ala, Gly, Pro, or similar residues may be employed. The substitutions may be homo-oligomers or hetero-oligomers. The number and types of residues which are substituted or added depend on the spacing necessary between essential contact points and certain functional attributes which are sought (e.g., hydrophobicity versus hydrophilicity). Increased binding affinity for an MHC molecule or T cell receptor may also be achieved by such substitutions, compared to the affinity of the parent peptide. In any event, such substitutions should employ amino acid residues or other molecular fragments chosen to avoid, for example, steric and charge interference which might disrupt binding.

[0042] Amino acid substitutions are typically of single residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final peptide. Substitutional variants are those in which at least one residue of a peptide has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Table 2 when it is desired to finely modulate the characteristics of the peptide.

TABLE 2

Original Residue	Exemplary Substitution
Ala	Ser
Arg	Lys, His
Asn	Gln
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Lys; Arg
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; His
Met	Leu; Ile
Phe	Tyr; Trp
Ser	Thr
Thr	Ser
Trp	Tyr; Phe
Tyr	Trp; Phe
Val	Ile; Leu

[0043] Substantial changes in function (e.g., affinity for MHC molecules or T cell receptors) are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the peptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in peptide properties will be those in which (a) hydrophilic residue, e.g. seryl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a residue having an electropositive side chain, e.g., lysl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (c) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

[0044] The peptides may also comprise isosteres of two or more residues in the immunogenic peptide. An isostere as defined here is a sequence of two or more residues that can be substituted for a second sequence because the steric conformation of the first sequence fits a binding site specific for the second sequence. The term specifically includes peptide backbone modifications well known to those skilled in the art. Such modifications include modifications of the amide nitrogen, the α-carbon, amide carbonyl, complete replacement of the amide bond, extensions, deletions or backbone crosslinks. See, generally, Spatola, Chemistry and Biochemistry of Amino Acids, peptides and Proteins, Vol. VII (Weinstein ed., 1983).

[0045] Modifications of peptides with various amino acid mimetics or unnatural amino acids are particularly useful in increasing the stability of the peptide in vivo. Stability can be assayed in a number of ways. For instance, peptidases and various biological media, such as human plasma and serum, have been used to test stability. See, e.g., Verhoef et al., Eur. J. Drug Metab. Pharmacokin. 11:291-302 (1986). Half life of the peptides of the present invention is conveniently determined using a 25% human serum (v/v) assay. The protocol is generally as follows. Pooled human serum (Type AB, non-heat inactivated) is delipidated by centrifugation before use. The serum is then diluted to 25% with RPMI tissue culture media and used to test peptide stability. At predetermined time intervals a small amount of reaction solution is removed and added to either 6% aqueous trichloracetic acid or ethanol. The cloudy reaction sample is cooled (4°C) for 15 minutes and then spun to pellet the precipitated serum proteins. The

presence of the peptides is then determined by reversed-phase HPLC using stability-specific chromatography conditions.

[0046] The peptides of the present invention or analogs thereof which have CTL stimulating activity may be modified to provide desired attributes other than improved serum half life. For instance, the ability of the peptides to induce CTL activity can be enhanced by linkage to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. Particularly preferred immunogenic peptides/T helper conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, e.g., Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues. Alternatively, the CTL peptide may be linked to the T helper peptide without a spacer. The immunogenic peptide may be linked to the T helper peptide either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the immunogenic peptide or the T helper peptide may be acylated. Exemplary T helper peptides include tetanus toxoid 830-843, influenza 307-319, malaria circumsporozoite 382-398 and 378-389.

[0047] In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes CTL. Lipids have been identified as agents capable of priming CTL in vivo against viral antigens. For example, palmitic acid residues can be attached to the alpha and epsilon amino groups of a Lys residue and then linked, e.g., via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be injected directly in a micellar form, incorporated into a liposome or emulsified in an adjuvant, e.g., incomplete Freund's adjuvant. In a preferred embodiment a particularly effective immunogen comprises palmitic acid attached to alpha and epsilon amino groups of Lys, which is attached via linkage, e.g., Ser-Ser, to the amino terminus of the immunogenic peptide.

[0048] As another example of lipid priming of CTL responses, E. coli lipoproteins, such as tripalmitoyl-S-glycerylcysteinlyseryl-serine (P_3CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide. See, Deres et al., Nature 342:561-564 (1989), incorporated herein by reference. Peptides of the invention can be coupled to P_3CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Further, as the induction of neutralizing antibodies can also be primed with P_3CSS conjugated to a peptide which displays an appropriate epitope, the two compositions can be combined to more effectively elicit both humoral and cell-mediated responses to infection.

[0049] In addition, additional amino acids can be added to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support, or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or the like, can be introduced at the C- or N-terminus of the peptide or oligopeptide. Modification at the C terminus in some cases may alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the natural sequence by being modified by terminal-NH_2 acylation, e.g., by alkanoyl (C_1-C_{20}) or thioglycolyl acetylation, terminal-carboxyl amidation, e.g., ammonia, methylamine, etc. In some instances these modifications may provide sites for linking to a support or other molecule.

[0050] The peptides of the invention can be prepared in a wide variety of ways. Because of their relatively short size, the peptides can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. See, for example, Stewart and Young, Solid Phase Peptide Synthesis, 2d. ed., Pierce Chemical Co. (1984), supra.

[0051] Alternatively, recombinant DNA technology may be employed wherein a nucleotide sequence which encodes an immunogenic peptide of interest is inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, New York (1982), which is incorporated herein by reference. Thus, fusion

proteins which comprise one or more peptide sequences of the invention can be used to present the appropriate T cell epitope.

[0052] As the coding sequence for peptides of the length contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci et al., J. Am. Chem. Soc. 103:3185 (1981), modification can be made simply by substituting the appropriate base(s) for those encoding the native peptide sequence. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly available in the art, and the vectors used to transform suitable hosts to produce the desired fusion protein. A number of such vectors and suitable host systems are now available. For expression of the fusion proteins, the coding sequence will be provided with operably linked start and stop codons, promoter and terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts. Of course, yeast or mammalian cell hosts may also be used, employing suitable vectors and control sequences.

[0053] The peptides of the present invention and pharmaceutical and vaccine compositions thereof are useful for administration to mammals, particularly humans, to treat and/or prevent viral infection and cancer. Examples of diseases which can be treated using the immunogenic peptides of the invention include prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV and condlyloma acuminatum.

[0054] For pharmaceutical compositions, the immunogenic peptides of the invention are administered to an individual already suffering from cancer or infected with the virus of interest. Those in the incubation phase or the acute phase of infection can be treated with the immunogenic peptides separately or in conjunction with other treatments, as appropriate. In therapeutic applications, compositions are administered to a patient in an amount sufficient to elicit an effective CTL response to the virus or tumor antigen and to cure or at least partially arrest symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, e.g., the peptide composition, the manner of administration, the stage and severity of the disease being treated, the

weight and general state of health of the patient, and the judgment of the prescribing physician, but generally range for the initial immunization (that is for therapeutic or prophylactic administration) from about 1.0 μg to about 5000 μg of peptide for a 70 kg patient, followed by boosting dosages of from about 1.0 μg to about 1000 μg of peptide pursuant to a boosting regimen over weeks to months depending upon the patient's response and condition by measuring specific CTL activity in the patient's blood. It must be kept in mind that the peptides and compositions of the present invention may generally be employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, in view of the minimization of extraneous substances and the relative nontoxic nature of the peptides, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions.

[0055] For therapeutic use, administration should begin at the first sign of viral infection or the detection or surgical removal of tumors or shortly after diagnosis in the case of acute infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In chronic infection, loading doses followed by boosting doses may be required.

[0056] Treatment of an infected individual with the compositions of the invention may hasten resolution of the infection in acutely infected individuals. For those individuals susceptible (or predisposed) to developing chronic infection the compositions are particularly useful in methods for preventing the evolution from acute to chronic infection. Where the susceptible individuals are identified prior to or during infection, for instance, as described herein, the composition can be targeted to them, minimizing need for administration to a larger population.

[0057] The peptide compositions can also be used for the treatment of chronic infection and to stimulate the immune system to eliminate virus-infected cells in carriers. It is important to provide an amount of immuno-potentiating peptide in a formulation and mode of administration sufficient to effectively stimulate a cytotoxic T cell response. Thus, for treatment of chronic infection, a representative dose is in the range of about 1.0 μg to about 5000 μg, preferably about 5 μg to 1000 μg for a 70 kg patient per dose. Immunizing doses followed by boosting doses at established intervals, e.g., from one to four weeks, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic infection,

administration should continue until at least clinical symptoms or laboratory tests indicate that the viral infection has been eliminated or substantially abated and for a period thereafter.

[0058] The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral or local administration. Preferably, the pharmaceutical compositions are administered parenterally, e.g., intravenously, subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, e.g., water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

[0059] The concentration of CTL stimulatory peptides of the invention in the pharmaceutical formulations can vary widely, i.e., from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in accordance with the particular mode of administration selected.

[0060] The peptides of the invention may also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or targeted selectively to infected cells, as well as increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to, e.g., a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the selected therapeutic/immunogenic peptide compositions.

Liposomes for use in the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., Ann. Rev. Biophys. Bioeng. 9:467 (1980), U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369, incorporated herein by reference.

[0061] For targeting to the immune cells, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, etc. in a dose which varies according to, inter alia, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

[0062] For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

[0063] For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, e.g., lecithin for intranasal delivery.

[0064] In another aspect the present invention is directed to vaccines which contain as an active ingredient an immunogenically effective amount of an immunogenic peptide as described herein. The peptide(s) may be introduced into a host, including humans, linked to its own carrier or as a homopolymer or heteropolymer of active peptide units. Such a polymer has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the virus or tumor cells. Useful carriers are well known in the art, and include, e.g., thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly(lysine: glutamic acid), influenza, hepatitis B virus core protein, hepatitis B virus recombinant vaccine and the like. The vaccines can also contain a physiologically tolerable (acceptable) diluent such as water, phosphate buffered saline, or saline, and further typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are materials well known in the art. And, as mentioned above, CTL responses can be primed by conjugating peptides of the invention to lipids, such as P_3CSS. Upon immunization with a peptide composition as described herein, via injection, aerosol, oral, transdermal or other route, the immune system of the host responds to the vaccine by producing large amounts of CTLs specific for the desired antigen, and the host becomes at least partially immune to later infection, or resistant to developing chronic infection.

[0065] Vaccine compositions containing the peptides of the invention are administered to a patient susceptible to or otherwise at risk of viral infection or cancer to elicit an immune response against the antigen and thus enhance the patient's own immune response capabilities. Such an amount is defined to be an "immunogenically effective dose." In this use, the precise amounts again depend on the patient's state of health and weight, the mode of administration, the nature of the formulation, etc., but generally range from about 1.0 μg to about 5000 μg per 70 kilogram patient, more commonly from about 10 μg to about 500 μg mg per 70 kg of body weight.

[0066] In some instances it may be desirable to combine the peptide vaccines of the invention with vaccines which induce neutralizing antibody responses to the virus of interest, particularly to viral envelope antigens.

[0067] For therapeutic or immunization purposes, the peptides of the invention can also be expressed by attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use

of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits a host CTL response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848, incorporated herein by reference. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (Nature 351:456-460 (1991)) which is incorporated herein by reference. A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., Salmonella typhi vectors and the like, will be apparent to those skilled in the art from the description herein.

[0068] Antigenic peptides may be used to elicit CTL ex vivo, as well. The resulting CTL, can be used to treat chronic infections (viral or bacterial) or tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a peptide vaccine approach of therapy. Ex vivo CTL responses to a particular pathogen (infectious agent or tumor antigen) are induced by incubating in tissue culture the patient's CTL precursor cells (CTLp) together with a source of antigen-presenting cells (APC) and the appropriate immunogenic peptide. After an appropriate incubation time (typically 1-4 weeks), in which the CTLp are activated and mature and expand into effector CTL, the cells are infused back into the patient, where they will destroy their specific target cell (an infected cell or a tumor cell).

[0069] The peptides may also find use as diagnostic reagents. For example, a peptide of the invention may be used to determine the susceptibility of a particular individual to a treatment regimen which employs the peptide or related peptides, and thus may be helpful in modifying an existing treatment protocol or in determining a prognosis for an affected individual. In addition, the peptides may also be used to predict which individuals will be at substantial risk for developing chronic infection.

[0070] The following examples are [is] offered by way of illustration, not by way of limitation.

Example 1

Class I antigen isolation

[0071] A flow diagram of an HLA-A antigen purification scheme is presented in Figure 1. Briefly, the cells bearing the appropriate allele were grown in large batches (6-8 liters yielding ~5 x 109 cells), harvested by centrifugation and washed. All cell lines were maintained in RPMI 1640 media (Sigma) supplemented with 10% fetal bovine serum (FBS) and antibiotics. For large-scale cultures, cells were grown in roller bottle culture in RPMI 1640 with 10% FBS or with 10% horse serum and antibiotics. Cells were harvested by centrifugation at 1500 RPM IEC-CRU5000 centrifuge with 259 rotor and washed three times with phosphate-buffered saline (PBS) (0.01 M P04, 0.154 M NaCl, pH 7.2).

[0072] Cells were pelleted and stored at -70°C or treated with detergent lysing solution to prepare detergent lysates. Cell lysates were prepared by the addition of stock detergent solution [1% NP-40 (Sigma) or Renex 30 (Accurate Chem. Sci. Corp., Westbury, NY 11590), 150 mM NaCl, 50 mM Tris, pH 8.0] to the cell pellets (previously counted) at a ratio of 50-100 x 10^6 cells per ml detergent solution. A cocktail of protease inhibitors was added to the premeasured volume of stock detergent solution immediately prior to the addition to the cell pellet. Addition of the protease inhibitor cocktail produced final concentrations of the following: phenylmethylsulfonyl fluoride (PMSF), 2 mM; aprotinin, 5 μg/ml; leupeptin, 10 μg/ml; pepstatin, 10 μg/ml; iodoacetamide, 100 μM; and EDTA, 3 ng/ml. Cell lysis was allowed to proceed at 4°C for 1 hour with periodic mixing. Routinely 5-10 x 10^9 cells were lysed in 50-100 ml of detergent solution. The lysate was clarified by centrifugation at 15,000 x g for 30 minutes at 4°C and subsequent passage of the supernatant fraction through a 0.2 μ filter unit (Nalgene).

[0073] The HLA-A antigen purification was achieved using affinity columns prepared with mAb-conjugated Sepharose beads. For antibody production, cells were grown in RPMI with 10% FBS in large tissue culture flasks (Corning 25160-225). Antibodies were purified from clarified tissue culture medium by ammonium sulfate fractionation followed by affinity chromatography on protein-A-Sepharose (Sigma). Briefly, saturated ammonium sulfate was added slowly with stirring to the tissue culture supernatant to 45% (volume to volume) overnight

at 4°C to precipitate the immunoglobulins. The precipitated proteins were harvested by centrifugation at 10,000 x g for 30 minutes. The precipitate was then dissolved in a minimum volume of PBS and transferred to dialysis tubing (Spectro/Por 2, Mol. wt. cutoff 12,000-14,000, Spectum Medical Ind.). Dialysis was against PBS (≥20 times the protein solution volume) with 4-6 changes of dialysis buffer over a 24-48 hour period at 4°C. The dialyzed protein solution was clarified by centrifugation (10,000 x g for 30 minutes) and the pH of the solution adjusted to pH 8.0 with 1N NaOH. Protein-A-Sepharose (Sigma) was hydrated according to the manufacturer's instructions, and a protein-A-Sepharose column was prepared. A column of 10 ml bed volume typically binds 50-100 mg of mouse IgG.

[0074] The protein sample was loaded onto the protein-A-Sepharose column using a peristaltic pump for large loading volumes or by gravity for smaller volumes (<100 ml). The column was washed with several volumes of PBS, and the eluate was monitored at A280 in a spectrophotometer until base line was reached. The bound antibody was eluted using 0.1 M citric acid at suitable pH (adjusted to the appropriate pH with 1N NaOH). For mouse IgG-1 pH 6.5 was used for IgG2a pH 4.5 was used and for IgG2b and IgG3 pH 3.0 was used. 2 M Tris base was used to neutralize the eluate. Fractions containing the antibody (monitored by A280) were pooled, dialyzed against PBS and further concentrated using an Amicon Stirred Cell system (Amicon Model 8050 with YM30 membrane). The anti-A2 mAb, BB7.2, was useful for affinity purification.

[0075] The HLA-A antigen was purified using affinity columns prepared with mAb-conjugated Sepharose beads. The affinity columns were prepared by incubating protein-A-Sepharose beads (Sigma) with affinity-purified mAb as described above. Five to 10 mg of mAb per ml of bead is the preferred ratio. The mAb bound beads were washed with borate buffer (borate buffer: 100 mM sodium tetraborate, 154 mM NaCl, pH 8.2) until the washes show A280 at based line. Dimethyl pimelimidate (20 mM) in 200 mM triethanolamine was added to covalently crosslink the bound mAb to the protein-A-Sepharose (Schneider et al., J. Biol. Chem. 257:10766 (1982). After incubation for 45 minutes at room temperature on a rotator, the excess crosslinking reagent was removed by washing the beads twice with 10-20 ml of 20 mM ethanolamine, pH 8.2. Between each one the slurry was placed on a rotator for 5 minutes at room temperature. The beads were washed with borate buffer and with PBS plus 0.02% sodium azide.

[0076] The cell lysate (5-10 x 10^9 cell equivalents) was then slowly passed over a 5-10 ml affinity column (flow rate of 0.1-0.25 ml per minute) to allow the binding of the antigen to the immobilized antibody. After the lysate was allowed to pass through the column, the column was washed sequentially with 20 column volumes of detergent stock solution plus 0.1% sodium dodecyl sulfate, 20 column volumes of 0.5 M NaCl, 20 mM Tris, pH 8.0, and 10 column volumes of 20 mM Tris, pH 8.0. The HLA-A antigen bound to the mAb was eluated with a basic buffer solution (50 mM diethylamine in water). As an alternative, acid solutions such as 0.15-0.25 M acetic acid were also used to elute the bound antigen. An aliquot of the eluate (1/50) was removed for protein quantification using either a colorimetric assay (BCA assay, Pierce) or by SDS-PAGE, or both. SDS-PAGE analysis was performed as described by Laemmli (Laemmli, U.K., Nature 227:680 (1970)) using known amounts of bovine serum albumin (Sigma) as a protein standard. Allele specific antibodies were used to purify the specific MHC molecule. In the case of HLA-A2, the mAb BB7.2 was used.

Example 2

Isolation and sequencing of naturally processed peptides

[0077] For the HLA-A preparations derived from the base (50 mM diethylamine) elution protocol, the eluate was immediately neutralized with 1 N acetic acid to pH 7.0-7.5. The neutralized eluate was concentrated to a volume of 1-2 ml in an Amicon stirred cell [Model 8050, with YM3 membranes (Amicon)]. Ten ml of ammonium acetate (0.01 M, pH 8.0) was added to the concentrator to remove the non-volatile salts, and the sample was concentrated to approximately 1 ml. A small sample (1/50) was removed for protein quantitation as described above. The remainder was recovered into a 15 ml polypropylene conical centrifuge tube (Falcon, 2097) (Becton Dickinson). Glacial acetic acid was added to obtain a final concentration of 10% acetic acid. The acidified sample was placed in a boiling water bath for 5 minutes to allow for the dissociation of the bound peptides. The sample was cooled on ice, returned to the concentrator and the filtrate was collected. Additional aliquots of 10% acetic acid (1-2 ml) were added to the concentrator, and this filtrate was pooled with the original filtrate. Finally, 1-2 ml of distilled water was added to the concentrator, and this filtrate was pooled as well.

[0078] The retentate contains the bulk of the HLA-A heavy chain and β_2-microglobulin, while the filtrate contains the naturally processed bound peptides and other components with molecular weights less than about 3000. The pooled filtrate material was lyophilized in order to concentrate the peptide fraction. The sample was then ready for further analysis.

[0079] For HPLC (high performance liquid chromatography) separation of the peptide fractions, the lyophilized sample was dissolved in 50 μl of distilled water, or into 0.1% trifluoracetic acid (TFA) (Applied Biosystems) in water and injected to a C18 reverse-phase narrow bore column (Beckman C18 Ultrasphere, 10 x 250 mm), using a gradient system described by Stone and Williams (Stone, K.L. and Williams K.R., in, Macromolecular Sequencing and Synthesis; Selected Methods and Applications, A.R. Liss, New York, 1988, pp. 7-24. Buffer A was 0.06% TFA in water (Burdick-Jackson) and buffer B was 0.052% TFA in 80% acetonitrile (Burdick-Jackson). The flow rate was 0.250 ml/minute with the following gradient: 060 min., 2-37.5% B; 60-95 min., 37.5-75% B; 95-105 min., 75-98% B. The Gilson narrow bore HPLC configuration is particularly useful for this purpose, although other configurations work equally well.

[0080] A large number of peaks were detected by absorbance at 214 nm, many of which appear to be of low abundance. Whether a given peak represents a single peptide or a peptide mixture was not determined. Pooled fractions were then sequenced to determine motifs specific for each allele as described below.

[0081] Pooled peptide fractions, prepared as described above were analyzed by automated Edman sequencing using the Applied Biosystems Model 477A automated sequencer. The sequencing method is based on the technique developed by Pehr Edman in the 1950s for the sequential degradation of proteins and peptides to determine the sequence of the constituent amino acids.

[0082] The protein or peptide to be sequenced was held by a 12-mm diameter porous glass fiber filter disk in a heated, argon-purged reaction chamber. The filter was generally pretreated with BioBrene Plus™ and then cycled through one or more repetitions of the Edman reaction to reduce contaminants and improve the efficiency of subsequent sample sequencing. Following the pre-treatment of the filter, a solution of the sample protein or peptide (10 pmol-5 nmol range) was loaded onto the glass filter and dried. Thus, the sample was left embedded in the film of the

pre-treated disk. Covalent attachment of the sample to the filter was usually not necessary because the Edman chemistry utilized relatively apolar solvents, in which proteins and peptides are poorly soluble.

[0083] Briefly, the Edman degradation reaction has three steps: coupling, cleavage, and conversion. In coupling step, phenylisothiocyanate (PITC) is added. The PITC reacts quantitatively with the free amino-terminal amino acid of the protein to form the phenylthiocarbamyl-protein in a basic environment. After a period of time for the coupling step, the excess chemicals are extracted and the highly volatile organic acid, trifluoroacetic acid, TFA, is used to cleave the PITC-coupled amino acid residue from the amino terminus of the protein yielding the anilinothiazolinone (ATZ) derivative of the amino acid. The remaining protein/peptide is left with a new amino terminus and is ready for the next Edman cycle. The ATZ amino acid is extracted and transferred to a conversion flask, where upon addition of 25% TFA in water, the ATZ amino acid is converted to the more stable phenylthiohydantoin (PTH) amino acid that can be identified and quantified following automatic injection into the Model 120 PTH Analyzer which uses a microbore C-18 reverse-phase HPLC column for the analysis.

[0084] In the present procedures, peptide mixtures were loaded onto the glass filters. Thus a single amino acid sequence usually does not result. Rather, mixtures of amino acids in different yield are found. When the particular residue is conserved among the peptides being sequenced, increased yield for that amino acid is observed.

Example 3

Definition of an A2.1 specific motif

[0085] In one case, pooled peptide fractions prepared as described in Example 2 above were obtained from HLA-A2.1 homozygous cell lines, for example, JY. The pooled fractions were HPLC fractions corresponding to 7% to 45% CH_3CN. For this class I molecule, this region of the chromatogram was most abundant in peptides. Data from independent experiments were averaged as described below.

[0086] The amino acid sequence analyses from four independent experiments were analyzed and the results are shown in Table 3. For each position except the first, the data were analyzed by modifying the method described by Falk et al., supra, to allow for comparison of experiments

from different HLA types. This modified procedure yielded quantitative yet standardized values while allowing the averaging of data from different experiments involving the same HLA type.

[0087] The raw sequenator data was converted to a simple matrix of 10 rows (each representing one Edman degradation cycle) and 16 columns (each representing one of the twenty amino acids; W, C, R and H were eliminated for technical reasons. The data corresponding to the first row (first cycle) was not considered further because, this cycle is usually heavily contaminated by free amino acids.). The values of each row were summed to yield a total pmoles value for that particular cycle. For each row, values for each amino acid were then divided by the corresponding total yield value, to determine what fraction of the total signal is attributable to each amino acid at each cycle. By doing so, an "Absolute Frequency" table was generated. This absolute frequency table allows correction for the declining yields of each cycle.

TABLE 3

A2.1: POOL SEQUENCING FREQUENCY

	pos. 1	pos. 2	pos. 3	pos. 4	pos. 5	pos. 6	pos. 7	pos. 8	pos. 9	pos. 10
A	-	0.65	1.25	0.85	0.95	0.77	1.21	1.16	1.15	1.25
G	-	0.84	0.98	1.29	1.22	0.89	0.78	1.05	0.98	1.48
D	-	0.84	1.11	[illegible]	1.03	0.83	0.82	0.84	0.82	1.19
E	-	0.38	0.59	[illegible]	1.10	0.82	1.05	1.45	0.87	0.88
R	-	-	-	-	-	-	-	-	-	-
H	-	-	-	-	-	-	-	-	-	-
K	-	0.63	0.65	0.89	[illegible]	1.09	0.89	1.35	0.82	0.87
L	-	[illegible]	1.11	0.45	0.57	1.00	0.89	0.59	0.92	0.77
V	-	0.78	0.69	0.60	0.79	1.38	1.24	0.84	[illegible]	1.27
I	-	1.06	1.20	0.53	0.93	1.49	1.15	0.76	0.88	0.54
M	-	[illegible]	[illegible]	0.62	0.71	0.68	0.88	0.54	0.73	0.22
Y	-	0.28	1.41	0.65	1.32	0.78	1.34	1.21	1.00	0.79
F	-	0.76	1.46	0.69	1.16	1.00	1.07	1.09	0.78	0.73
W	-	-	-	-	-	-	-	-	-	-
Q	-	0.60	0.84	0.92	0.95	0.90	1.16	[illegible]	1.00	1.00
N	-	0.39	0.76	1.17	1.28	1.08	1.07	1.28	0.96	0.42
S	-	1.13	[illegible]	1.33	0.87	0.77	0.71	0.92	0.77	0.58
T	-	0.82	0.90	0.94	0.95	1.21	1.07	[illegible]	0.71	0.57
C	-	-	-	-	-	-	-	-	-	-
P	-	0.54	0.78	1.44	1.15	1.09	1.30	0.87	0.81	1.01

* [illegible] significant increase

[0088] Starting from the absolute frequency table, a "relative frequency" table was then generated to allow comparisons among different amino acids. To do so the data from each column was summed, and then averaged. Then, each value was divided next by the average column value to obtain relative frequency values. These values quantitate, in a standardized manner, increases and decreases per cycle, for each of the different sixteen amino acid types. Tables generated from data from different experiments can thus be added together to generate average relative frequency values (and their standard deviations). All standard deviations can then be averaged, to estimate a standard deviation value applicable to the samples from each table. Any particular value exceeding 1.00 by more than two standard deviations is considered to correspond to a significant increase.

Example 4

Quantitative Binding Assays

[0089] Using isolated MHC molecules prepared as described in Example 2, above, quantitative binding assays were performed. Briefly, indicated amounts of MHC as isolated above were incubated in 0.05% NP40-PBS with ~5 nM of radiolabeled peptides in the presence of 1-3 μM β_2M and a cocktail of protease inhibitors (final concentrations 1 mM PMSF, 1.3 mM 1.10 Phenanthroline, 73 μM Pepstatin A, 8 mM EDTA, 200 μM N-α-p-tosyl-L-Lysine Chloromethyl ketone). After various times, free and bound peptides were separated by TSK 2000 gel filtration, as described previously in A. Sette et al., J. Immunol. 148:844 (1992), which is incorporated herein by reference. Peptides were labeled by the use of the Chloramine T method Buus et al., Science 235:1352 (1987), which is incorporated herein by reference.

[0090] The HBc 18-27 peptide HLA binding peptide was radiolabeled and offered (5-10 nM) to 1 μM purified HLA A2.1. After two days at 23°C in presence of a cocktail of protease inhibitors and 1-3 μM purified human β_2M, the percent of MHC class I bound radioactivity was measured by size exclusion chromatography, as previously described for class II peptide binding assays in Sette et al., in Seminars in Immunology, Vol. 3, Gefter, ed. (W.B. Saunders, Philadelphia, 1991), pp 195-202, which is incorporated herein by reference. Using this protocol, high binding (95%) was detected in all cases in the presence of purified HLA A2.1 molecules.

[0091] To explore the specificity of binding, we determined whether the binding was inhabitable by excess unlabeled peptide, and if so, what the 50% inhibitory concentration (IC50%) might be. The rationale for this experiment was threefold. First, such an experiment is crucial in order to demonstrate specificity. Second, a sensitive inhibition assay is the most viable alternative for a high throughput quantitative binding assay. Third, inhibition data subjected to Scatchard analysis can give quantitative estimates of the equilibrium constant (K) of interaction and the fraction of receptor molecules capable of binding ligand (% occupancy). For instance, in analysis of an inhibition curve for the interaction of the peptide HBc 18-27 with A2.1, the IC50% was determined to be 25 nM. Further experiments were conducted to obtain Scatchard plots. For HBc 18-27/A2.1, six different experiments using six independent MHC preparations yielded a K_D Of $15.5 \pm 9.9 \times 10^{-9}$ M and occupancy values of 6.2% (±1.4).

[0092] Several reports have demonstrated that class I molecules, unlike class II, are highly selective with regard to the size of the peptide epitope that they recognize. The optimal size varies between 8 and 10 residues for different peptides and different class I molecules, although MHC binding peptides as long as 13 residues have been identified. To verify the stringent size requirement, a series of N- and C-terminal truncation/extension analogs of the peptide HBc 18-27 were synthesized and tested for A2.1 binding. Previous studies had demonstrated that the optimal size for CTL recognition of this peptide was the 10-mer HBcl8-27 (Sette et al. supra). It was found that removal or addition of a residue at the C terminus of the molecule resulted in a 30 to 100-fold decrease in binding capacity. Further removal or addition of another residue completely obliterated binding. Similarly, at the N-terminus of the molecule, removal or deletion of one residue from the optimal HBc 18-27 peptide completely abrogated A2.1 binding.

[0093] Throughout this disclosure, results have been expressed in terms of IC50's. Given the conditions in which our assays are run (i.e., limiting MHC and labeled peptide concentrations), these values approximate K_D values. It should be noted that IC50 values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents used (e.g., Class I preparation, etc.). For example, excessive concentrations of MHC will increase the apparent measured IC50 of a given ligand.

[0094] An alternative way of expressing the binding data, to avoid these uncertainties, is as a relative value to a reference peptide. The reference peptide is included in every assay. As a

particular assay becomes more, or less, sensitive, the IC50's of the peptides tested may change somewhat. However, the binding relative to the reference peptide will not change. For example, in an assay run under conditions such that the IC50 of the reference peptide increases 10-fold, all IC50 values will also shift approximately ten-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good, intermediate, weak, or negative binder should be based on it's IC50, relative to the IC50 of the standard peptide.

[0095] The reference peptide for the HLA-A2.1 assays described herein is referred to as 941.01 having a sequence of FLPSDYFPSV. An average IC50 of 5 (nM) was observed under the assay conditions utilized.

[0096] If the IC50 of the standard peptide measured in a particular assay is different from that reported in the table, then it should be understood that the threshold values used to determine good, intermediate, weak, and negative binders should be modified by a corresponding factor. For example, if in an A2.1 binding assay, the IC50 of the A2.1 standard (941.01) were to be measured as 8 nM instead of 5 nM, then a peptide ligand would be called a good binder only if it had an IC50 of less than 80 nM (i.e., 8nM x 0.1), instead of the usual cut-off value of 50 nM.

Example 5

HLA-A2.1 Binding Motif and Algorithm

[0097] The structural requirements for peptide binding to A2.1 have been defined for both, 9-mer and 10-mer peptides. Two approaches have been used. The first approach referred to as the "poly-A approach" uses a panel of single amino acid substitutions of a 9-mer prototype poly-A binder (ALAKAAAAV) that is tested for A2.1 binding using the methods of Example 4 above to examine the degree of degeneracy of the anchor-positions and the possible influence of non-anchor positions on A2.1 binding.

[0098] The second approach, the "Motif-Library approach", uses a large library of peptides selected from sequences of potential target molecules of viral and tumor origin and tested for A2.1 binding using the methods in Example 4 above. The frequencies by which different amino-acids occurred at each position in good binders and non-binders were analyzed to further define the role of non-anchor positions in 9-mers and 10-mers.

A2.1 binding of peptide 9-mers

[0099] Poly A Approach A poly-A 9-mer peptide, containing the A2.1 motif L (Leu) in position 2 and V (Val) in position 9 was chosen as a prototype binder. A K (Lys) residue was included in position 4 to increase solubility. A panel of 91 single amino-acid substitution analogues of the prototype parental 9-mer was synthesized and tested for A2.1 binding (Table 4). Shaded areas mark analogs with a greater than 10-fold reduction in binding capacity relative to the parental peptide. A reduction in binding greater than 100-fold is indicated a dash.

[00100] Anchor-Positions 2 and 9 in poly-A Analogs The effect of single-amino-acid substitutions at the anchor positions 2 and 9 was examined first. Most substitutions in these positions had profound detrimental effects on binding capacity, thus confirming their role for binding. More specifically, in position 2 only L and M bound within a 10-fold range ("preferred residues"). Residues with similar characteristics, such as I, V, A, and T were tolerated, but bound 10 to 100-fold less strongly than the parental peptide. All the remaining substitutions (residues S, N, D, F, C, K, G, and P) were not tolerated and decreased binding by more than 100-fold. Comparably stringent requirements were observed for position 9, where V, L and I were preferred and A and M are tolerated, while the residues T, C, N, F, and Y virtually abolished binding. According to this set of peptides, an optimal 2-9 motif could be defined with L, M in position 2 and V, I, or L in position 9.

TABLE 4

A2.1 binding of analogs of a motif-containing poly A peptide

	pos. 1 A	pos. 2 L	pos. 3 A	pos. 4 K	pos. 5 A	pos. 6 A	pos. 7 A	pos. 8 A	pos. 9 V
A	1.00	[illegible]	1.00		1.00	1.00	1.00	1.00	[illegible]
C	0.46				0.63	0.12		0.57	
D			0.93	0.74	0.51	[illegible]			
E	[illegible]		0.68	1.53	0.62	0.15	0.28	0.26	
R						[illegible]			
H								0.24	
K	0.54	[illegible]		1.00	0.39		0.50	0.11	
L		1.00	0.46		0.99		0.76	0.90	0.11
V	0.47	[illegible]	0.15	1.12		0.44	0.49	0.30	1.00
I	0.41	[illegible]				1.12			0.18
M		0.43	0.66						[illegible]
Y	0.75		0.62		0.94	0.41	1.40	0.43	
F	1.10		0.95			1.76		0.49	
W									
Q					0.32		0.19	0.41	
N			0.34		1.24		0.97	0.31	
S	0.44		0.37	0.97					
T	0.28	[illegible]		0.98			0.28	0.37	
C				1.53		0.84			
P			0.25	1.07		0.84	0.83	0.55	

[illegible] Ratio ≤ 0.1

• Ratio ≤ 0.01

[00101] Non-Anchor Positions 1 and 3-8 in poly-A Analogs All non-anchor positions were more permissive to different substitutions than the anchor-positions 2 and 9, i.e. most residues were tolerated. Significant decreases in binding were observed for some substitutions in distinct positions. More specifically, in position 1 a negative charge (residues D and E) or a P greatly reduced the binding capacity. Most substitutions were tolerated in position 3 with the exception of the residue K. Significant decreases were also seen in position 6 upon introduction of either a negative charge (D, E) or a positively charged residue (R). A summary of these effects by different single amino acid substitutions is given in Table 5.

TABLE 5

Summary A2.1 Poly-A

AA position	(+)	(+/-)	(-)
1	FAYKVGSIT		EDP
2	LM	VITA	SNDFCKGP
3	AFDEMYLSNPV	K	
4	CEVPATSD		
5	NALYGEDKQ		
6	FIAPCVYEG	DR	
7	YANLPVETQ		
8	ALGPFYQTNVEHK		
9	VIL	AM	TCNFY
	Ratio>0.1	Ratio 0.01-0.1	Ratio<0.01

[00102] The Motif-Library Approach To further evaluate the importance of non-anchor positions for binding, peptides of potential target molecules of viral and tumor origin were scanned for the presence of sequences containing optimal 2-9 anchor motifs. A set of 161 peptides (appendix I) containing a L or M in position 2 and a V, L or I in position 9 was selected, synthesized and tested for binding (see Table 17). Only 11.8% of these peptides bind with high affinity (ratio $\geq$0.10; 22.4% were intermediate binders (ratio $\geq$0.1). As many as 36% were weak binders (ratio <0.01 - 0.0001), and 31% were non-binders (ratio<0.0001). The high number of non-binders containing optimal anchor-motifs indicates that in this set of peptides positions other than the 2-9 anchors influence A2.1 binding capacity. Appendix 1 sets forth all of the peptides having the 2-9 motif used for this analysis and the binding data for those peptides.

[00103] To define the influence of non-anchor positions more specifically, the frequency of occurrence of each amino acid in each of the non-anchor positions was calculated for the good and intermediate binders on one hand and non-binders on the other hand. Amino acids of similar chemical characteristic were grouped together (Table 6). Weak binders were not considered for the following analysis. The frequency of occurrence of each amino acid in each of the non-anchor positions was calculated for the good binders and non-binders (Table 6).

[00104] Several striking trends become apparent. For example in position 1, only 3.6% of the A2.1 binders and as much as 35% of the non-binders carried a negative charge (residues D and E). This observation correlates well with previous findings in the set of poly-A analogs, where a D or E substitution greatly affected binding. Similarly, the residue P was 8 times more frequent in non-binders than in good binders. Conversely, the frequencies of aromatic residues (Y, F, W) were greatly increased in A2.1 binders as compared to non-binders.

Table 6

A2.1 9-mer PEPTIDES

NUMBER OF PEPTIDES	161	
GOOD BINDERS	19	11.8%
INTERMEDIATE BINDERS	36	22.4%
WEAK BINDERS	58	36.0%
NON-BINDERS	48	29.8%

	1+	1-	2+	2-	3+	3-	4+	4-	5+	5-	6+	6-	7+	7-	8+	8-	9+	9-
A	5.5	2.1	0.0	0.0	3.6	4.2	5.5	8.3	5.5	8.3	5.5	6.3	9.1	2.1	3.6	12.5	0.0	0.0
G	7.3	2.1	0.0	0.0	3.6	8.3	9.1	8.3	9.1	8.3	10.9	8.3	5.5	12.5	3.6	8.3	0.0	0.0
D,E	3.6	35.4	0.0	0.0	0.0	12.5	10.9	16.7	3.6	12.5	5.5	8.3	1.8	16.7	9.1	10.4	0.0	0.0
R,H,K	12.7	4.2	0.0	0.0	3.6	16.7	16.4	16.7	9.1	10.4	1.8	20.8	0.0	10.4	16.4	12.5	0.0	0.0
L,V,I,M	38.2	12.5	100.0	100.0	34.5	18.8	9.1	16.7	25.5	29.2	30.9	22.9	30.9	25.0	32.7	18.8	100.0	100.0
Y,F,W	14.5	2.1	0.0	0.0	21.8	4.2	7.3	8.3	18.2	2.1	16.4	8.3	14.5	8.3	5.5	2.1	0.0	0.0
Q,N	7.3	14.6	0.0	0.0	5.5	14.6	12.7	10.4	9.1	10.4	10.9	10.4	5.5	8.3	5.5	16.7	0.0	0.0
S,T,C	9.1	12.5	0.0	0.0	20.0	10.4	20.0	4.2	14.5	16.7	14.5	12.5	14.5	12.5	20.0	18.8	0.0	0.0
P	1.8	14.6	0.0	0.0	7.3	10.4	9.1	12.5	5.5	2.1	3.6	2.1	18.2	8.3	3.6	0.0	0.0	0.0
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

[00105] Following this approach, amino acids of similar structural characteristics were grouped together. Then, the frequency of each amino acid group in each position was calculated for binders versus non-binders. Finally, the frequency in the binders group was divided by the frequency in the non-binders to obtain a "frequency ratio". (Table 7) This ratio indicates whether a given amino-acid or group of residues occurs in a given position preferentially in good binders (ratio >1) or in non-binders (ratio <1).

TABLE 7

A2.1 9-mer PEPTIDES

NUMBER OF PEPTIDES	161	
GOOD BINDERS	19	11.8%
INTERMEDIATE BINDERS	36	22.4%
WEAK BINDERS	58	36.0%
NON-BINDERS	48	29.8%

	pos. 1 ratio	pos. 2 ratio	pos. 3 ratio	pos. 4 ratio	pos. 5 ratio	pos. 6 ratio	pos. 7 ratio	pos. 8 ratio	pos. 9 ratio
A	2.6	NA	0.9	0.9	0.7	0.9	4.4	0.3	NA
G	3.5	NA	0.4	1.1	1.1	1.3	0.4	0.4	NA
D,E	0.1	NA	0.0	0.7	0.3	0.7	0.1	0.9	NA
R,H,K	3.1	NA	0.2	1.0	0.9	0.1	0.0	1.3	NA
L,V,I,M	3.1	1.0	1.8	0.5	0.9	1.3	1.2	1.7	1.0
Y,F,W	7.0	NA	5.2	0.9	8.7	2.0	2.3	2.6	NA
Q,N	0.5	NA	0.4	1.2	0.9	1.0	0.7	0.3	NA
S,T,C	0.7	NA	1.9	4.8	0.9	1.2	1.2	1.1	NA
P	0.1	NA	0.7	0.7	2.6	1.7	2.9	+++	NA

+++ indicatos that there were no negative binders

[00106] Different Residues Influence A2.1 Binding In order to analyze the most striking influences of certain residues on A2.1 binding, a threshold level was set for the ratios described in Table 7. Residues showing a more than 4-fold greater frequency in good binders were regarded as preferred residues (+). Residues showing a 4-fold lower frequency in A2.1 binders than in non-binders were regarded as disfavored residues (-). Following this approach, residues showing the most prominent positive or negative effects on binding are listed in Table 8.

[00107] This table identifies the amino acid groups which influence binding most significantly in each of the non-anchor positions. In general, the most negative effects were observed with charged amino acids. In position 1, negatively charged amino acids were not observed in good binders, i.e., those amino acids were negative binding residues at position 1. The opposite was true for position 6 where only basic amino acids were detrimental for binding i.e., were negative binding residues. Moreover, both acidic and basic amino acids were not observed in A2.1 binders in positions 3 and 7. A greater than 4-fold increased frequency of non-binders was found when P was in position 1.

TABLE 8

Summary of A2.1 Motif-Library , 9-mers

AA position	(+)	(-)
1	(YFW)	P, (DE)
2	Anchor	
3	(YFW)	(DE), (RKH)
4	(STC)	
5	(YFW)	
6		(RKH)
7	A	(RKH), (DE)
8		
9	Anchor	

(+) = Ratio ≥ 4-fold (-) = Ratio ≤ 0.25

[00108] Aromatic residues were in general favored in several of the non-anchor positions, particularly in positions 1, 3, and 5. Small residues like S, T, and C were favored in position 4 and A was favored in position 7.

[00109] An Improved A2.1 9-mer Motif The data described above was used to derive a stringent A2.1 motif. This motif is based in significant part on the effects of the non-anchor positions 1 and 3-8. The uneven distribution of amino acids at different positions is reflective of specific dominant negative binding effects of certain residues, mainly charged ones, on binding affinity. A series of rules were derived to identify appropriate anchor residues in positions 2 and 9 and negative binding residues at positions 1 and 3-8 to enable selection of a high affinity binding immunogenic peptide. These rules are summarized in Table 9.

[00110] To validate the motif defined above and shown in Table 9 published sequences of peptides that have been naturally processed and presented by A2.1 molecules were analyzed (Table 10). Only 9-merpeptides containing the 2-9 anchor residues were considered.

[00111] When the frequencies of these peptides were analyzed, it was found that in general they followed the rules summarized in Table 9. More specifically, neither acidic amino acids nor P were found in position 1. Only one acidic amino acid and no basic amino acids were found in position 3. Positions 6 and 7 showed no charged residues. Acidic amino acids, however, were frequently found in position 8, where they are tolerated, according to our definition of the A2.1 motif. The analysis of the sequences of naturally processed peptides therefore reveals that >90% of the peptides followed the defined rules for a complete motif.

[00112] Thus the data confirms a role of positions other than the anchor positions 2 and 9 for A2.1 binding. Most of the deleterious effects on binding are induced by charged amino acids in non-anchor positions, i.e. negative binding residues occupying positions 1, 3, 6 or 7.

TABLE 9

A2.1 Motif for 9-mer peptides

AA Position	(+)	(-)
1		acidic amino acids and P
2	Anchor: L, M, I, V, A, T	
3		acidic and basic amino acids
4		
5		
6		basic amino acids
7		acidic and basic amino acids
8		
9	Anchor: V, L, I, A, M	

TABLE 10

A2.1 natur ally proce ssed pep tides

POSITIONS									
1	2	3	4	5	6	7	8	9	A2.1 binding
A	L	X	G	G	X	V	N	V	ND
L	L	D	V	P	T	A	A	V	ND
G	X	V	P	F	X	V	S	V	0.41
S	L	L	P	A	I	V	E	L	0.19
S	X	X	V	R	A	X	E	V	ND
Y	M	N	G	T	M	S	Q	V	ND
K	X	N	E	P	V	X	X	X	ND
Y	L	L	P	A	I	V	H	I	0.26
A	X	W	G	F	F	P	V	X	ND
T	L	W	V	D	P	Y	E	V	0.23
G	X	V	P	F	X	V	S	V	0.41

A2.1 Binding of Peptide 10-mers

[00113] The "Motif-Library" Approach Previous data clearly indicated that 10-mers can also bind to HLA molecules even if with a somewhat lower affinity than 9-mers. For this reason we expanded our analysis to 10-mer peptides.

[00114] Therefore, a "Motif-Library" set of 170 peptide 10-mers containing optimal motif-combinations was selected from known target molecule sequences of viral and tumor origin and analyzed as described above for 9-mers. In this set we found 5.9% good binders, 17.1% intermediate binders, 41.2% weak binders and 35.9% non-binders. The actual sequences, origin and binding capacities of this set of peptides are included as Appendix 2. This set of 10-mers was used to determine a) the rules for 10-mer peptide binding to A2.1, b) the similarities or differences to rules defined for 9-mers, and c) if an insertion point can be identified that would allow for a superimposable common motif for 9-mers and 10-mers.

[00115] Amino-acid frequencies and frequency ratios for the various amino-acid groups for each position were generated for 10-mer peptides as described above for 9-mer peptides and are also shown in Tables 11 and 12, respectively for grouped residues.

[00116] A summary of preferred versus disfavored residues and of the rules derived for the 10-mers in a manner analogous to that used for 9-mers, is also listed in Tables 13 and 14, respectively.

[00117] When the frequency-ratios of different amino-acid groups in binders and non-binders at different positions were analyzed and compared to the corresponding ratios for the 9-mers, both striking similarities and significant differences emerged (Table 15). At the N-terminus and the C-termini of 9-mers and 10-mers, similarities predominate. In position 1 for example, in 10-mers again the P residue and acidic amino acids were not tolerated. In addition at position 1 in 10-mers aromatic residues were frequently observed in A2.1 binders. In position 3, acidic amino acids were frequently associated with poor binding capacity in both 9-mers and 10-mers. Interestingly, however, while in position 3 aromatic residues were preferred in 9-mers, aliphatic residues (L, V, I, M) were preferred in 10-mers.

Table 11

A2.1 10-mer Peptides

NUMBER OF PEPTIDES	170	
GOOD BINDERS	10	5.9%
INTERMEDIATE BINDERS	29	17.1%
WEAK BINDERS	70	41.2%
NON-BINDERS	61	35.9%

	1+	1-	2+	2-	3+	3-	4+	4-	5+	5-	6+	6-	7+	7-	8+	8-	9+	9-	10+	10-
A	2.6	0.0	0.0	0.0	10.3	3.3	2.6	11.5	5.1	3.3	7.7	13.1	10.3	8.2	7.7	4.9	2.6	4.9	0.0	0.0
G	7.7	9.8	0.0	0.0	7.7	16.4	15.4	3.3	5.1	6.6	10.3	1.6	17.9	6.6	7.7	11.5	7.7	9.8	0.0	0.0
D,E	0.0	23.0	0.0	0.0	2.6	16.4	7.7	13.1	2.6	9.8	10.3	0.8	5.1	16.4	0.0	16.4	5.1	13.1	0.0	0.0
R,H,K	7.7	6.6	0.0	0.0	5.1	16.4	2.6	16.0	10.3	14.8	7.7	19.7	2.6	14.8	0.0	29.5	2.6	16.4	0.0	0.0
L,V,I,M	48.7	16.4	100.0	100.0	33.3	3.3	23.1	23.0	30.8	24.6	30.8	14.8	25.6	18.0	23.1	4.9	12.8	16.4	100.0	100.0
Y,F,W	12.8	0.0	0.0	0.0	12.8	4.9	15.4	4.9	17.9	4.9	7.7	13.1	12.8	6.2	23.1	1.6	20.5	9.8	0.0	0.0
Q,N	10.3	9.8	0.0	0.0	7.7	8.2	7.7	9.8	7.7	9.8	2.6	3.3	5.1	8.2	2.6	6.6	7.7	11.5	0.0	0.0
S,T,C	10.3	11.5	0.0	0.0	15.4	18.0	12.8	11.5	20.5	19.7	17.9	19.7	17.9	13.1	20.5	16.4	33.3	11.5	0.0	0.0
P	0.0	23.0	0.0	0.0	5.1	13.1	12.8	4.9	0.0	6.6	5.1	4.9	2.6	6.6	15.4	8.2	7.7	6.6	0.0	0.0
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 12

A2.1 10-mer Peptides

NUMBER OF PEPTIDES	170	
GOOD BINDERS	10	5.9%
INTERMEDIATE BINDERS	29	17.1%
WEAK BINDERS	70	41.2%
NON-BINDERS	61	35.9%

	pos. 1 ratio	pos. 2 ratio	pos. 3 ratio	pos. 4 ratio	pos. 5 ratio	pos. 6 ratio	pos. 7 ratio	pos. 8 ratio	pos. 9 ratio	pos. 10 ratio
A	+++	NA	3.1	0.2	1.8	0.6	1.3	1.6	0.5	NA
G	0.8	NA	0.5	4.7	0.8	6.3	2.7	0.7	0.8	NA
D,E	0.0	NA	0.2	0.6	0.3	1.0	0.3	0.0	0.4	NA
R,H,K	1.2	NA	0.3	0.1	0.7	0.4	0.2	0.0	0.2	NA
L,V,I,M	3.0	1.0	10.2	1.0	1.3	2.1	1.4	4.7	0.8	1.0
Y,F,W	+++	NA	2.6	3.1	3.6	0.6	1.6	14.1	2.1	NA
Q,N	1.0	NA	0.9	0.8	0.8	0.8	0.6	0.4	0.7	NA
S,T,C	0.9	NA	0.9	1.1	1.0	0.9	1.4	1.3	2.9	NA
P	0.0	NA	0.4	2.6	0.0	1.0	0.4	1.9	1.2	NA

+++ Indicates that there were no negative binders.

TABLE 13

Summary of A2.1 Motif-Library , 10-mers

AA position	(+)	(-)
1	(YFW), A	(DE), P
2	Anchor	
3	(LVIM)	(DE)
4	G	A,(RKH)
5		P
6	G	
7		(RKH)
8	(YFW), (LVIM)	(DE), (RKH)
9		(RKH)
10	Anchor	

(+) = Ratio ≥ 4-fold (-) = Ratio ≤ 0.25

TABLE 14

A2.1 Motif for 10-mer peptides

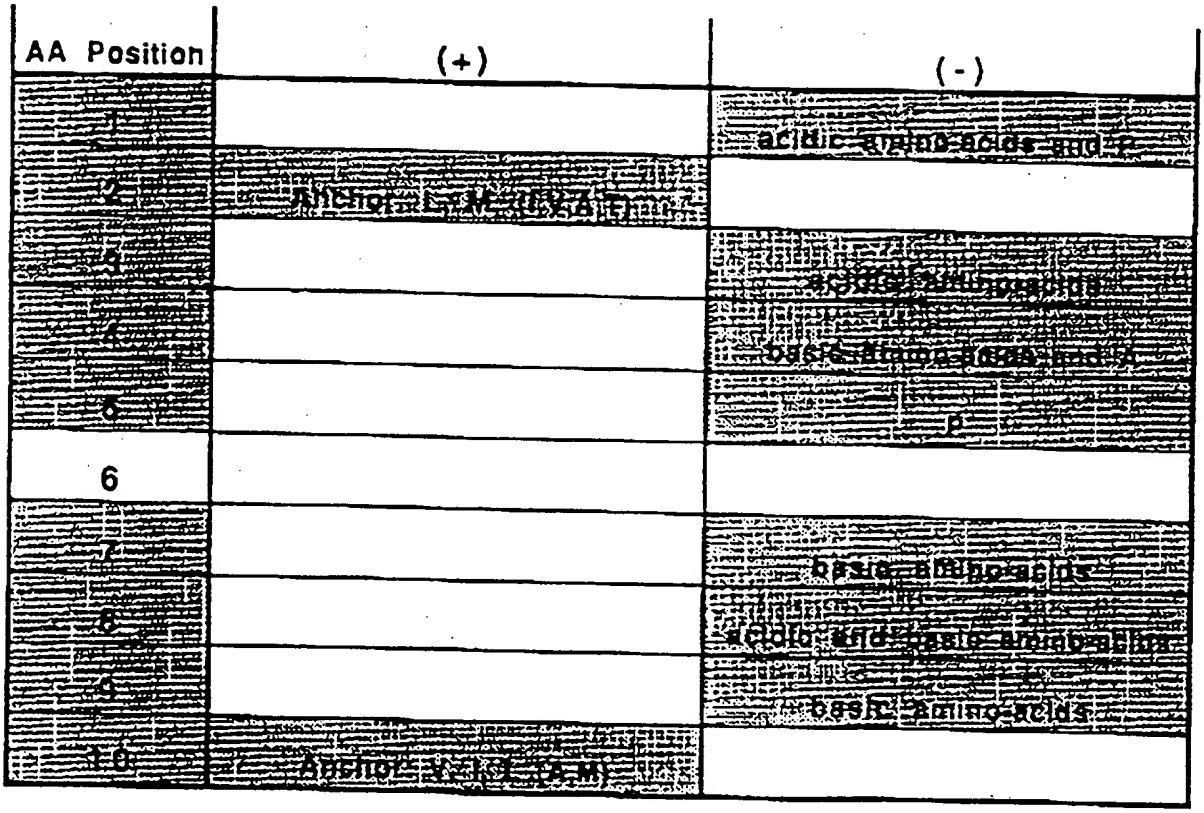

TABLE 15

Comparison of A2.1 binding of 9-mers and of 10-mers

AA Position	9-mers (+)	10-mers (+)
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	(YFW)	(LVIM)
4	(STC)	G
5	(YFW)	
6		G
7	A	
8		(YFW) (LVIM)
9	[illegible]	
10	-	[illegible]

AA Position	9-mers (-)	10-mers (-)
1	[illegible]	[illegible]
2		
3	[illegible]	[illegible]
4		A,(RKH)
5		P
6	[illegible]	
7	[illegible]	[illegible]
8		[illegible]
9	[illegible]	(RKH)
10	-	[illegible]

[00118] At the C-terminus of the peptides, basic amino acids are not favored in position 7, and both acidic and basic amino acids are not favored in position 8 for 10-mers. This is in striking agreement with the observation that the same pattern was found in 9-mers at positions 6 and 7. Interestingly, again the favored residues differ between two peptide's sizes. Aromatic (Y, F, W) or aliphatic (L, V, I, M) residues were preferred in 10-mers at position 8, while the A residue was preferred by 9-mers in the corresponding position 7.

[00119] By contrast, in the center of the peptide no similarities of frequency preferences were observed at positions 4, 5, and 6 in 10-mers and positions 4 and 5 in the 9-mers.

[00120] Most interestingly, among the residues most favored in the center of the tested peptides were G in position 4 and 6, P in position 5 was not observed in binders. All of these residues are known to dramatically influence the overall secondary structure of peptides, and in particular would be predicted to strongly influence the propensity of a 10-mer to adopt a "kinked" or "bulged" conformation.

[00121] Charged residues are predominantly deleterious for binding and are frequently observed in non-binders of 9-mers and 10-mers.

[00122] However, favored residues are different for 9-mers and 10-mers. Glycine is favored while Proline is disfavored in the center of 10-mer peptides but this is not the case for 9-mers.

[00123] These data establish the existence of an "insertion area" spanning two positions (4, 5) in 9-mers and 3 positions (4, 5, 6) in 10-mers. This insertion area is a more permissive region where few residue similarities are observed between the 9-mer and 10-mer antigenic peptides. Furthermore, in addition to the highly conserved anchor positions 2 and 9, there are "anchor areas" for unfavored residues in positions 1 and 3 at the N-terminus for both 9-mer and 10-mer and positions 7-10 or 6-9 at the C-terminus for 10-mers and 9-mers, respectively.

Example 6

Algorithm to Predict Binding of 9-mer Peptides to HLA-A2.1

[00124] Within the population of potential A2.1 binding peptides identified by the 2,9 motif, as shown in the previous example, only a few peptides are actually good or intermediate

binders and thus potentially immunogenic. It is apparent from the data previously described that the residues present in positions other than 2 and 9 can influence, often profoundly, the binding affinity of a peptide. For example, acidic residues at position 1 for A2.1 peptides do not appear to be tolerated. Therefore, a more exact predictor of binding could be generated by taking into account the effects of different residues at each position of a peptide sequence, in addition to positions 2 and 9.

[00125] More specifically, we have utilized the data bank obtained during the screening of our collection of A2.1 motif containing 9-mer peptides to develop an algorithm which assigns a score for each amino acid, at each position along a peptide. The score for each residue is taken as the ratio of the frequency of that residue in good and intermediate binders to the frequency of occurrence of that residue in non-binders.

[00126] In the present "Grouped Ratio" algorithm residues have been grouped by similarity. This avoids the problem encountered with some rare residues, such as tryptophan, where there are too few occurrences to obtain a statistically significant ratio. Table 16 is a listing of scores obtained by grouping for each of the twenty amino acids by position for 9-mer peptides containing perfect 2/9 motifs. A peptide is scored in the "Grouped Ratio" algorithm as a product of the scores of each of its residues. In the case of positions other than 2 and 9, the scores have been derived using a set of peptides which contain only preferred residues in positions 2 and 9. To enable us to extend our "Grouped Ratio" algorithm to peptides which may have residues other than the preferred ones at 2 and 9, scores for 2 and 9 have been derived from a set of peptides which are single amino acid substitutions at positions 2 and 9. Figure 2 shows a scattergram of the log of relative binding plotted against "Grouped Ratio" algorithm score for our collection of 9-mer peptides from the previous example.

TABLE 16

	1	2	3	4	5	6	7	8	9
A	2.6	0.03	0.87	0.87	0.65	0.87	4.4	0.29	0.16
C	0.73	0.01	1.9	4.8	0.87	1.2	1.2	1.1	0.01
D	0.10	0.01	0.10	0.65	0.29	0.65	0.11	0.87	0.01
E	0.10	0.01	0.10	0.65	0.29	0.65	0.11	0.87	0.01
F	7.0	0.01	5.2	0.87	8.7	2.0	2.3	2.6	0.01
G	3.5	.0.01	0.44	1.1	1.1	1.3	0.44	0.44	0.01
H	3.1	0.01	0.22	1.0	0.87	0.09	0.10	1.3	0.01
I	3.1	0.14	1.8	0.55	0.87	1.4	1.2	1.8	0.40
K	3.1	0.01	0.22	1.0	0.87	0.09	0.10	1.3	0.01
L	3.1	1.00	1.8	0.55	0.87	1.4	1.2	1.8	0.09
M	3.1	2.00	1.8	0.55	0.87	1.4	1.2	1.8	0.06
N	0.50	0.01	0.37	1.2	0.87	1.1	0.65	0.33	0.01
P	0.12	0.01	0.70	0.73	2.6	1.8	2.9	0.10	0.01
Q	0.50	0.01	0.37	1.2	0.87	1.1	0.65	0.33	0.01
R	3.1	0.01	0.22	1.0	0.87	0.09	0.10	1.3	0.01
S	0.73	0.01	1.9	4.8	0.87	1.2	1.2	1.1	0.01
T	0.73	0.01	1.9	4.8	0.87	1.2	1.2	1.1	0.01
V	3.1	0.08	1.8	0.55	0.87	1.4	1.2	1.8	1.00
W	7.0	0.01	5.2	0.87	8.7	2.0	2.3	2.6	0.01
Y	7.0	0.01	5.2	0.87	8.7	2.0	2.3	2.6	0.01

[00127] The present "Grouped Ratio" algorithm can be used to predict a population of peptides with the highest occurrence of good binders. If one were to rely, for example, solely on a 2(L,M) and 9(L,I, and V) motif for predicting A2.1 binding 9-mer peptides, it would have been predicted that all 161 peptides in our database would be good binders. In fact, as has already been described, only 12% of these peptides would be described as good binders and only 22% as intermediate binders; 66% of the peptides predicted by such a 2,9 motif are either weak or non-binding peptides. In contrast, using the "Grouped Ratio" algorithm described above, and selecting a score of 1.0 as threshold, 20 peptides were selected. Of this set, 50% are good binders, and 45% are intermediate, while only 5% are weak and 0% are non-binders (Table 17).

[00128] The present example of an algorithm has used the ratio of binders/non-binders to measure the impact of a particular residue at each position of a peptide. It is immediately apparent to one of ordinary skill that there are alternative ways of creating a similar algorithm.

[00129] An algorithm using the average binding affinity of all the peptides with a certain amino acid (or amino acid type) at a certain position has the advantage of including all of the peptides in the analysis, and not just good/intermediate binders and non-binders. Moreover, it gives a more quantitative measure of affinity than the simpler "Grouped Ratio" algorithm. We have created such an algorithm by calculating for each amino acid, by position, the average log of binding when that particular residue occurs in our set of 161 2,9 motif containing peptides. These values are shown in Table 18. The algorithm score for a peptide is then taken as the sum of the scores by position for each residues. Figure 3 shows a scattergram of the log of relative binding against the average "Log of Binding" algorithm score. Table 17 shows the ability of the two algorithms to predict peptide binding at various levels, as a function of the cut-off score used. The ability of a 2,9 motif to predict binding in the same peptide set is also shown for reference purposes. It is clear from this comparison that both algorithms of this invention have a greater ability to predict populations with higher frequencies of good binders than a 2,9 motif alone. Differences between the "Grouped Ratio" algorithm and the "Log of Binding" algorithm are small in the set of peptides analyzed here, but do suggest that the "Log of Binding" algorithm is a better, if only slightly, predictor than the "Grouped Ratio" algorithm.

Example 7

Use of an Algorithm to Predict Binding of 10-mer Peptides to HLA-A2.1

[00130] Using the methods described in the proceeding example, an analogous set of algorithms has been developed for predicting the binding of 10-mer peptides. Table 19 shows the scores used in a "Grouped Ratio" algorithm, and Table 20 shows the "Log of Binding" algorithm scores, for 10-mer peptides. Table 21 shows a comparison of the application of different algorithmic methods to select binding peptides. Figures 4 and 5 show, respectively, scattergrams of a set of 10-mer peptides containing preferred residues in positions 2 and 10 as scored by the "Grouped Ratio" and "Log of Binding" algorithms.

Example 8

Binding of A2.1 Algorithm Predicted Peptides

[00131] The results of Examples 6 and 7 indicate that an algorithm can be used to select peptides that bind to HLA-A2.1 sufficiently to have a high probability of being immunogenic.

[00132] To test this result, we tested our algorithm on a large (over 1300) non-redundant, independent set of peptides derived from various sources. After scoring this set with our algorithm, we selected 41 peptides (Table 22) for synthesis, and tested them for A2.1 binding. This set of peptides was comprised of 21 peptides with high algorithm scores, and 20 peptides with low algorithm scores.

Table 17

Criteria	Cut-off	Good Binders	Intermediate Binders	Weak Binders	Negative Binders	Totals
2,9 motif		19 (12%)	36 (22%)	58 (36%)	48 (30%)	161 (100%)
Grouped Ratio Algorithm	1.5	5 (83%)	1 (17%)	0 (0%)	0 (0%)	6 (100%)
	1.25	8 (67%)	4 (33%)	0 (0%)	0 (0%)	12 (100%)
	1	10 (50%)	9 (45%)	1 (5%)	0 (0%)	20 (100%)
	0.5	12 (32%)	17 (46%)	7 (19%)	1 (3%)	37 (100%)
	0	12 (23%)	26 (49%)	12 (23%)	3 (6%)	53 (100%)
	-1	17 (18%)	35 (37%)	33 (35%)	10 (11%)	95 (100%)
	-2	19 (15%)	36 (29%)	50 (40%)	21 (17%)	126 (100%)
	-3	19 (13%)	36 (24%)	56 (38%)	38 (26%)	149 (100%)
	no cut	19 (12%)	36 (22%)	58 (36%)	48 (30%)	161 (100%)
Log of Binding Algorithm	-19	5 (100%)	0 (0%)	0 (0%)	0 (0%)	5 (100%)
	-20	8 (73%)	3 (27%)	0 (0%)	0 (0%)	11 (100%)
	-21	15 (43%)	15 (43%)	5 (14%)	0 (0%)	35 (100%)
	-22	17 (26%)	27 (41%)	21 (32%)	1 (2%)	66 (100%)
	-23	18 (19%)	35 (37%)	34 (36%)	7 (7%)	94 (100%)
	-24	19 (16%)	36 (30%)	47 (39%)	17 (14%)	119 (100%)
	-25	19 (14%)	36 (26%)	55 (39%)	30 (21%)	140 (100%)
	no cut	19 (12%)	36 (22%)	58 (36%)	48 (30%)	161 (100%)

TABLE 18

	1	2	3	4	5	6	7	8	9
A	-2.38	-3.22	-2.80	-2.68	-2.89	-2.70	-2.35	-3.07	-2.49
C	-2.94	-4.00	-2.58	-1.96	-3.29	-2.22	-2.97	-2.37	-4.00
D	-3.69	-4.00	-3.46	-2.71	-2.26	-2.63	-3.61	-3.03	-4.00
E	-3.64	-4.00	-3.51	-2.65	-3.39	-3.41	-3.21	-2.63	-4.00
F	-1.89	-4.00	-2.35	-2.50	-1.34	-2.43	-2.18	-1.71	-4.00
G	-2.32	-4.00	-3.04	-2.63	-2.56	-2.30	-3.13	-2.96	-4.00
H	-2.67	-4.00	-2.58	-2.58	-2.05	-3.32	-3.13	-2.16	-4.00
I	-1.65	-2.55	-2.80	-3.44	-2.74	-2.79	-2.20	-2.69	-2.10
K	-2.51	-4.00	-3.65	-2.93	-3.34	-3.77	-3.13	-3.27	-4.00
L	-2.32	-1.70	-2.02	-2.49	-2.71	-2.63	-2.62	-2.01	-2.74
M	-0.39	-1.39	-1.79	-3.07	-3.43	-1.38	-1.33	-0.97	-2.96
N	-3.12	-4.00	-3.52	-2.22	-2.36	-2.30	-3.14	-3.31	-4.00
P	-3.61	-4.00	-2.97	-2.64	-2.42	-2.31	-1.83	-2.42	-4.00
Q	-2.76	-4.00	-2.81	-2.63	-3.06	-2.84	-2.12	-3.05	-4.00
R	-1.92	-4.00	-3.41	-2.61	-3.05	-3.76	-3.43	-3.02	-4.00
S	-2.39	-3.52	-2.04	-2.12	-2.83	-3.04	-2.73	-2.02	-4.00
T	-2.92	-4.00	-2.60	-2.48	-2.17	-2.58	-2.67	-3.14	-3.70
V	-2.44	-2.64	-2.68	-3.29	-2.49	-2.24	-2.68	-2.83	-1.70
W	-0.14	-4.00	-1.01	-2.94	-1.63	-2.77	-2.85	-2.13	-4.00
X	-1.99	-2.13	-2.41	-2.97	-2.72	-2.70	-2.41	-2.35	-2.42
Y	-1.46	-4.00	-1.67	-2.70	-1.92	-2.39	-1.35	-3.37	-4.00

TABLE 19

	1	2	3	4	5	6	7	8	9	10
A	3.00	0.01	3.10	0.20	1.60	0.60	1.30	1.60	0.50	0.01
C	0.90	0.01	0.90	1.10	1.00	0.90	1.40	1.30	2.90	0.01
D	0.01	0.01	0.20	0.60	0.30	1.00	0.30	0.01	0.40	0.01
E	0.01	0.01	0.20	0.60	0.30	1.00	0.30	0.01	0.40	0.01
F	3.00	0.01	2.60	3.10	3.60	0.60	1.60	14.1	2.10	0.01
G	0.80	0.01	0.50	4.70	0.80	6.30	2.70	0.70	0.80	0.01
H	1.20	0.01	0.30	0.10	0.70	0.40	0.20	0.01	0.20	0.01
I	3.00	0.50	10.2	1.00	1.30	2.10	1.40	4.70	0.80	1.00
K	1.20	0.01	0.30	0.10	0.70	0.40	0.20	0.01	0.20	0.01
L	3.00	1.10	10.2	1.00	1.30	2.10	1.40	4.70	0.80	0.50
M	3.00	0.60	10.2	1.00	1.30	2.10	1.40	4.70	0.80	0.01
N	1.00	0.01	0.90	0.80	0.80	0.80	0.60	0.40	0.70	0.01
P	0.00	0.01	0.40	2.60	0.01	1.00	0.40	1.90	1.20	0.01
Q	1.00	0.01	0.90	0.80	0.80	0.80	0.60	0.40	0.70	0.01
R	1.20	0.01	0.30	0.10	0.70	0.40	0.20	0.01	0.20	0.01
S	0.90	0.01	0.90	1.10	1.00	0.90	1.40	1.30	2.90	0.01
T	0.90	0.01	0.90	1.10	1.00	0.90	1.40	1.30	2.90	0.01
V	3.00	0.10	10.2	1.00	1.30	2.10	1.40	4.70	0.80	2.30
W	3.00	0.01	2.60	3.10	3.60	0.60	1.60	14.1	2.10	0.01
Y	3.00	0.01	2.60	3.10	3.60	0.60	1.60	14.1	2.10	0.01

TABLE 20

	1	2	3	4	5	6	7	8	9	10
A	-2.40	-4.00	-2.54	-3.42	-3.07	-3.30	-2.98	-2.69	-3.29	-4.00
C	-3.64	-4.00	-2.47	-2.48	-1.78	-3.94	-1.28	-3.10	-2.43	-4.00
D	-3.65	-4.00	-2.76	-3.26	-2.76	-3.03	-3.43	-3.68	-3.63	-4.00
E	-3.92	-4.00	-3.63	-3.34	-3.73	-2.82	-3.54	-3.71	-2.95	-4.00
F	-1.52	-4.00	-1.96	-3.03	-2.01	-3.11	-2.67	-1.61	-2.43	-4.00
G	-2.91	-4.00	-3.40	-2.63	-2.98	-2.45	-2.52	-3.18	-3.03	-4.00
H	-3.61	-4.00	-3.10	-3.03	-2.33	-2.99	-3.70	-3.55	-4.00	-4.00
I	-2.26	-4.00	-2.82	-3.05	-2.38	-2.61	-2.38	-3.34	-3.18	-1.47
K	-2.53	-4.00	-3.65	-3.42	-3.14	-3.58	-3.50	-3.53	-4.00	-4.00
L	-2.00	-2.93	-2.21	-2.48	-2.88	-2.53	-2.57	-1.83	-3.23	-3.20
M	-2.41	-3.11	-2.00	-3.33	-3.70	-2.56	-3.27	-2.25	-3.00	-4.00
N	-3.21	-4.00	-3.09	-2.61	-2.93	-2.89	-3.52	-3.01	-2.88	-4.00
P	-3.90	-4.00	-3.21	-2.27	-3.72	-3.06	-3.35	-2.58	-2.94	-4.00
Q	-2.92	-4.00	-2.97	-4.00	-2.98	-3.46	-2.20	-3.23	-3.45	-4.00
R	-3.01	-4.00	-3.44	-3.50	-3.23	-3.32	-3.72	-3.59	-2.97	-4.00
S	-2.47	-4.00	-3.17	-3.11	-3.23	-2.64	-3.19	-2.79	-2.26	-4.00
T	-3.59	-4.00	-3.07	-2.88	-2.89	-3.16	-2.43	-3.11	-2.58	-4.00
V	-2.97	-4.00	-2.46	-3.14	-3.27	-2.53	-3.14	-3.02	-2.90	-2.61
W	-2.10	-4.00	-2.72	-1.79	-2.65	-1.92	-1.80	-2.24	-2.11	-4.00
Y	-2.37	-4.00	-2.42	-2.85	-3.03	-3.76	-2.82	-2.34	-2.74	-4.00

Table 21

Criteria	Cut-off	Good Binders	Intermediate Binders	Weak Binders	Negative Binders	Totals
2,10 motif		10 (6%)	29 (17%)	70 (41%)	61 (36%)	170 (100%)
Grouped Ratio Algorithm	4	1 (100%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)
	3	1 (25%)	2 (50%)	1 (25%)	0 (0%)	4 (100%)
	2	6 (24%)	13 (52%)	6 (24%)	0 (0%)	25 (100%)
	1	10 (21%)	21 (45%)	16 (34%)	0 (0%)	47 (100%)
	0	10 (15%)	28 (42%)	26 (39%)	2 (3%)	66 (100%)
	-1	10 (11%)	29 (32%)	42 (46%)	11 (12%)	92 (100%)
	-2	10 (9%)	29 (25%)	54 (47%)	23 (20%)	116 (100%)
	-3	10 (7%)	29 (22%)	63 (47%)	32 (24%)	134 (100%)
	no cut	10 (6%)	29 (17%)	70 (41%)	61 (36%)	170 (100%)
Log of Binding Algorithm	-24	2 (50%)	2 (50%)	0 (0%)	0 (0%)	4 (100%)
	-25	5 (56%)	3 (33%)	1 (11%)	0 (0%)	9 (100%)
	-26	7 (47%)	5 (33%)	3 (20%)	0 (0%)	15 (100%)
	-27	10 (32%)	9 (29%)	12 (39%)	0 (0%)	31 (100%)
	-28	10 (17%)	19 (33%)	29 (50%)	0 (0%)	58 (100%)
	-29	10 (12%)	25 (30%)	48 (58%)	0 (0%)	83 (100%)
	-30	10 (10%)	29 (28%)	59 (57%)	5 (5%)	103 (100%)
	-31	10 (8%)	28 (22%)	66 (51%)	24 (19%)	128 (100%)
	-32	10 (7%)	29 (19%)	70 (47%)	40 (27%)	149 (100%)
	no cut	10 (6%)	29 (17%)	70 (41%)	61 (36%)	170 (100%)

[00133] The binding data and categorization profile are shown in Tables 22 and 23 respectively. The correlation between binding and algorithm score was 0.69. The striking difference between peptides with high algorithm scores is immediately apparent from Table 22, and those with low algorithm scores. Respectively, 76% of the high scorers and none of the low scorers were either good or intermediate binders. This data demonstrates the utility of the algorithm of this invention.

Table 22

SEQUENCE	SOURCE	A2.1 Binding	Algorithm Score
MMWFVVLTV	CMV	0.76	346
YLLLYFSPV	CMV	0.75	312
YLYRLNFCL	CMV	0.72	169
FMWTYLVTL	CMV	0.68	336
LLWWITLLL	CMV	0.49	356
GLWCVLFFV	CMV	0.47	1989
LMIRGVLEV	CMV	0.45	296
LLLCRLPFL	CMV	0.42	1356
RLLTSLFFL	HSV	0.34	859
LLLYYDYSL	HSV	0.28	390
AMSRNLFRV	CMV	0.15	1746
AMLTAGVEV	CMV	0.089	411
RLQPNVPLV	CMV	0.048	392
VLARTFTPV	CMV	0.044	196
RLLRGLRL	CMV	0.037	494
WMWFPSVLL	CMV	0.036	362
YLCCGITLL	CMV	0.021	1043
DMLGRVFFV	HSV	0.011	1422
ALGRYQQLV	CMV	0.0089	184
LMPPPVAEL	CMV	0.0066	418
LMCRYTPRL	CMV	0.0055	414
RLTWRLTWL	CMV	0.0052	250
AMPRRVLHV	CMV	0.0014	628
ALLLVLALL	CMV	0.0014	535
AMSGTGTTL	CMV	0.0005	602
MLNVMKEAV	CMV	0.0099	0.00031
TMELMRTV	CMV	0.0029	0.0013
TLAAMHSKL	HSV	0.0008	0.0019
TLNIVRQHV	CMV	0.0005	0.00021
ELSPREFL	HSV	0.0002	0.0020
FLRVQQKAL	HSV	0.0002	0.00099
ELQMMQCWV	CMV	0.0001	0.0020
QLNAMKPQL	MT	0.0001	0.0017
GLRQLKGAL	CMV	0.0001	0.0010
TLRMSSKAV	HSV	0.0001	0.00085
SLRKRELL	CMV	0	0.00041
DLKQMERVV	CMV	0	0.00026
PLRVTPSDL	CMV	0	0.0019
QLDYEKQVL	CMV	0	0.0012
WLKLLRDAL	CMV	0	0.0012
PMEAVRHPL	CMV	0	0.0011
ELKQTRVNL	CMV	0	0.00053
NLEVIHDAL	CMV	0	0.00050
ELKKVKSVL	HSV	0	0.00033
PLAYERDKL	CMV	0	0.00017

Table 23

Set	Good Binders	Intermediate Binders	Weak Binders	Negative Binders	Totals
HI Scorers	11 (52.4%)	5 (23.8%)	5 (23.8%)	0 (0.0%)	21 (100%)
Low Scorers	0 (0.0%)	0 (0.0%)	10 (50.0%)	10 (50.0%)	20 (100%)
Totals	11 (26.8%)	5 (12.2%)	15 (36.6%)	10 (24.4%)	41 (100%)

Example 9

Ex vivo induction of Cytotoxic T Lymphocytes (CTL)

[00134] Peripheral blood mononuclear cells (PBMC) are isolated from an HLA-typed patient by either venipuncture or apheresis (depending upon the initial amount of CTLp required), and purified by gradient centrifugation using Ficoll-Paque (Pharmacia). Typically, one can obtain one million PBMC for every ml of peripheral blood, or alternatively, a typical apheresis procedure can yield up to a total of 1-10 X 10^{10} PBMC.

[00135] The isolated and purified PBMC are co-cultured with an appropriate number of antigen presenting cell (APC), previously incubated ("pulsed") with an appropriate amount of synthetic peptide (containing the HLA binding motif and the sequence of the antigen in question). PBMC are usually incubated at 1-2 X 10^6 cells/ml in culture medium such as RPMI-1640 (with autologous serum or plasma) or the serum-free medium AIM-V (Gibco).

[00136] APC are usually used at concentrations ranging from lX10^4 to 2XlO^5 cells/ml, depending on the type of cell used. Possible sources of APC include: 1) autologous dendritic cells (DC), which are isolated from PBMC and purified as described (Inaba, et al., J. Exp. Med. 166:182 (1987)); and 2) mutant and genetically engineered mammalian cells that express "empty" HLA molecules (which are syngeneic [genetically identical] to the patient's allelic HLA form), such as the, mouse RMA-S cell line or the human T2 cell line. APC containing empty HLA molecules are known to be potent inducers of CTL responses, possibly because the peptide can associate more readily with empty MHC molecules than with MHC molecules which are occupied by other peptides (DeBruijn, et al., Eur. J. Immunol. 21:2963-2970 (1991)).

[00137] In those cases when the APC used are not autologous, the cells will have to be gamma irradiated with an appropriate dose (using, e.g., radioactive cesium or cobalt) to prevent their proliferation both ex vivo, and when the cells are re-introduced into the patients.

[00138] The mixture cultures, containing PBMC, APC and peptide are kept in an appropriate culture vessel such as plastic T-flasks, gas-permeable plastic bags, or roller bottles, at 37° centigrade in a humid air/CO_2 incubator. After the activation phase of the culture, which usually occurs during the first 3-5 days, the resulting effector CTL can be further expanded, by the addition of recombinant DNA-derived growth factors such as interleukin-2 (IL-2),

interleukin-4 (IL-4), or interleukin-7 (IL-7) to the cultures. An expansion culture can be kept for an additional 5 to 12 days, depending on the numbers of effector CTL required for a particular patient. In addition, expansion cultures may be performed using hollow fiber artificial capillary systems (Cellco), where larger numbers of cells (up to $1X10^{11}$) can be maintained.

[00139] Before the cells are infused into the patient, they are tested for activity, viability, toxicity and sterility. The cytotoxic activity of the resulting CTL can be determined by a standard ^{51}Cr-release assay (Biddison, W.E. 1991, Current Protocols in Immunology, p7.17.1-7.17.5, Ed. J. Coligan et al., J. Wiley and Sons, New York), using target cells that express the appropriate HLA molecule, in the presence and absence of the immunogenic peptide. Viability is determined by the exclusion of trypan blue dye by live cells. Cells are tested for the presence of endotoxin by conventional techniques. Finally, the presence of bacterial or fungal contamination is determined by appropriate microbiological methods (chocolate agar, etc.). Once the cells pass all quality control and safety tests, they are washed and placed in the appropriate infusion solution (Ringer/glucose lactate) and infused intravenously into the patient.

Example 10

Assays for CTL Activity

[00140] 1. Peptide synthesis: Peptide syntheses were carried out by sequential coupling of N-α-Fmoc-protected amino acids on an Applied Biosystems (Foster City, CA) 430A peptide synthesizer using standard Fmoc coupling cycles (software version 1.40). All amino acids, reagents, and resins were obtained from Applied Biosystems or Bachem. Solvents were obtained from Burdick & Jackson. Solid-phase synthesis was started from an appropriately substituted Fmoc-amino acid-Sasrin resin. The loading of the starting resin was 0.5-0.7 mmol/g polystyrene, and 0.1 or 0.25 meq were used in each synthesis. A typical reaction cycle proceeded as follows: 1) The N-terminal Fmoc group was removed with 25% piperidine in dimethylformamide (DMF) for 5 minutes, followed by another treatment with 25% piperdine in DMF for 15 minutes. The resin was washed 5 times with DMF. An N-methylpyrolidone (NMP) solution of a 4 to 10 fold excess of a pre-formed 1-hydroxybenzotriazole ester of the appropriate Fmoc-amino acid was added to the resin and the mixture was allowed to react for 30-90 min. The resin was washed with DMF in preparation for the next elongation cycle. The fully

protected, resin bound peptide was subjected to a piperidine cycle to remove the terminal Fmoc group. The product was washed with dichloromethane and dried. The resin was then treated with trifluoroacetic acid in the presence of appropriate scavengers [e.g. 5% (v/v) water] for 60 minutes at 20°C. After evaporation of excess trifluoroacetic acid, the crude peptide was washed with dimethyl ether, dissolved in water and lyophilized. The peptides wee purified to >95% homogeneity by reverse-phase HPLC using H_20/CH_3CN gradients containing 0.2% TFA modifier on a Vydac, 300Å pore-size, C-18 preparative column. The purity of the synthetic peptides was assayed on an analytical reverse-phase column, and their composition ascertained by amino acid analysis and/or sequencing. Peptides were routinely dissolved in DMSO at the concentration of 20 mg/ml.

[00141] 2. Media: RPMI-1640 containing 10% fetal calf serum (FCS) 2 mM Glutamine, 50 μg/ml Gentamicin and $5X10^{-5}$M 2-mercaptoethanol served as culture medium and will be referred to as R10 medium.

[00142] RPMI-1640 containing 25 mM Hepes buffer and supplemented with 2% FCS was used as cell washing medium.

[00143] 3. Rat Concanavalin A supernatant: The spleen cells obtained from Lewis rats (Sprague-Dawley) were resuspended at a concentration of $5lO^6$ cells/ml in R10 medium supplemented with 5 μg/ml of ConA in 75 cm2 tissue culture flasks. After 48 hr at 37°C, the supernatants were collected, supplemented with 1% α-methyl-D-mannoside and filter sterilized (.45 μm filter). Aliquots were stored frozen at -20°C.

[00144] 4. LPS-activated lymphoblasts: Murine splenocytes were resuspended at a concentration of $1\text{-}1.5xlO^6$/ml in R10 medium supplemented with 25 μg/ml LPS and 7 μg/ml dextran sulfate in 75 cm^2 tissue culture flasks. After 72 hours at 37°C, the lymphoblasts were collected for use by centrifugation.

[00145] 5. Peptide coating of lymphoblasts: Coating of the LPS activated lymphoblasts was achieved by incubating $3OxlO^6$ lymphoblasts with 100 μg of peptide in 1 ml of R10 medium for 1 hr at 37°C. Cells were then washed once and resuspended in R10 medium at the desired concentration for use in in vitro CTL activation.

[00146] 6. Peptide coating of Jurkat A2/K^b cells: Peptide coating was achieved by incubating $10xlO^6$ irradiated 20,000 rads) Jurkat A2.1/K^b cells with 20 μg of peptide in 1 ml of

R10 medium for 1 hour at 37°C. Cells were washed three times and resuspended at the required concentration in R10 medium.

[00147] 7. In Vitro CTL activation: One to four weeks after priming spleen cells (5 x 10^6 cells/well or $3OxlO^6$ cells/T25 flask) were concultured at 37°C with syngeneic, irradiated (3,000 rads), peptide coated lymphoblasts ($2xlO^6$ cells/well or $1OX10^6$ cells/T25 flask) in R10 medium to give a final volume of 2 ml in 24-well plates or 10 ml in T25 flasks.

[00148] 8. Restimulation of effector cells: Seven to ten days after the initial in vitro activation, described in paragraph 8 above, a portion of the effector cells were restimulated with irradiated (20,000 rads), peptide-coated Jurkat $A2/K^b$ cells ($0.2xlO^6$ cells/well) in the presence of $3xlO^6$ "feeder cells"/well (C57Bl/6 irradiated spleen cells) in RIO medium supplemented with 5% rat ConA supernatant to help provide all of the cytokines needed for optimal effector cell growth.

[00149] 9. Assay for cytotoxic activity: -Target cells ($3xlO^6$) were incubated at 37°C in the presence of 200 μl of sodium ^{51}Cr chromate. After 60 minutes, cells were washed three times and resuspended in R10 medium. Peptide 875.15 was added at the required concentration. For the assay, 10^4 ^{51}Cr-labeled target cells wee added to different concentrations of effector cells (final volume of 200 μl) in U-bottom 96-2311 plates. After a 6-hour incubation period at 37°C, 0.1 ml aliquots of supernatant were removed from each well and radioactivity was determined in a Micromedic automatic gamma counter. The percent specific lysis was determined by the formula: percent specific release = 100x(experimental release - spontaneous release)/(maximum release - spontaneous release). Where peptide titrations wee performed, the antigenicity of a given peptide (for comparison purposes) was expressed as the peptide concentration required to induce 40% specific ^{51}Cr release at a given E:T.

[00150] Transgenic mice were injected subcutaneously in the base of the tail with an incomplete Freund's adjuvant emulsion containing 50 nM of the putative CTL epitopes containing the A2.1 motifs, and 50 nM of the hepatitis B core T helper epitope, Cytel No. 875.23. Eight to 20 days later, animals were sacrificed and spleen cells were restimulated in vitro with syngeneic LPS lymphoblasts coated with the purative CTL epitope. A source of IL-2 (rat con A supernatant) was added at day 6 of the assay to a final concentration of 5% and CTL activity was measured on day 7. The capacity of these effector T cells to lyse peptide-coated target cells that

express the A2 KB molecule (Jurkat A2 KB) was measured as lytic units. The results are presented in Table 23.

[00151] The results of this experiment indicate that those peptides having a binding of at least 0.01 are capable of inducing CTL. All of the peptides in Appendices 1 and 2 having a binding of at least about 0.01 would be immunogenic.

TABLE 24

Binding and Immunogenicity
HBV Polymerase (ayw)

Peptide	Binding**	CTL Activity	Algorithm
1 2 3 4 5 6 7 8 9			
F L L S L G I H L	0.52	63	-20.8
G L Y S S T V P V	0.15	10	-21.9
H L Y S H P I I L	0.13	10	-21.1
W I L R G T S F V	0.018	-+	-20.9
N L S W L S L D V	0.013	6	-24.7
L L S S N L S W L	0.005	-	-21.7
N L Q S L T N L L	0.003	-	-23.9
H L L V G S S G L	0.002	-	-24.7
L L D D E A G P L	0.0002	-	-25.5
P L E E E L P R L	0.0001	-	-26.1
D L N L G N L N V	-*	-	-25.7
N L Y V S L L L L	-	-	-23.6
P L P I H T A E L	-	-	-25.04

*-=<0.0001
** Relative binding capacity compared to std with IC_{50} = 52mM
xxx Lytic units/10^6 cells; 1 lytic unit = the number of effector cells required to give 30% Cr^{51} release.
-,-+ no measurable cytotoxic activity.

Example 11[1]

[00152] Class I antigen isolation was carried out as described in the parent applications. Naturally processed peptides were then Isolated and sequenced as described there. An allele-specific motif and algorithms were determined and quantitative binding assays were carried.

[00153] Using the motifs identified above for HLA-A2.1 allele amino acid sequences from a tumor-related proteins, Melanoma Antigen-1,-2, and -3 (MAGE-1, -2, and -3), were analyzed for the presence of these motifs. Sequences for the target antigen are obtained from the GenBank data base (Release No. 71.0; 3/92). The identification of motifs is done using the "FINDPATTERNS" program (Devereux et al., Nucleic Acids Research 12:387-395 (1984)).

[00154] Other viral and tumor-related proteins can also be analyzed for the presence of these motifs. The amino acid sequence or the nucleotide sequence encoding products is obtained from the GenBank database in the cases of Human Papilloma Virus (HPV), Prostate Specific antigen (PSA), p53 oncogene, Epstein Barr Nuclear Antigen-1 (EBNA-1), and c-erb2 oncogene (also called HER-2/neu).

[00155] In the cases of Hepatitis B Virus (HBV), Hepatitis C Virus (HCV), and Human Immunodeficiency Virus (HIV) several strains/isolates exist and many sequences have been placed in GenBank.

[00156] For HBV, binding motifs are identified for the adr, adw and ayw types. In order to avoid replication of identical sequences, all of the adr motifs and only those motifs from adw and ayw that are not present in adr are added to the list of peptides.

[00157] In the case of HCV, a consensus sequence from residue 1 to residue 782 is derived from 9 viral isolates. Motifs are identified on those regions that have no or very little (one residue) variation between the 9 isolates. The sequences of residues 783 to 3010 from 5 viral isolates were also analyzed. Motifs common to all the isolates are identified and added to the peptide list.

[00158] Finally, a consensus sequence for HIV type 1 for North American viral isolates (10-12 viruses) was obtained from the Los Alamos National Laboratory database (May 1991 release) and analyzed in order to identify motifs that are constant throughout most viral isolates.

Motifs that bear a small degree of variation (one residue, in 2 forms) were also added to the peptide list.

[00159] Table 14 provides the results of searches of the following antigens cERB2, EBNA1, HBV, HCV, HIV, HPV, MAGE, p53, and PSA. Only peptides with binding affinity of at least 1% as compared to the standard peptide in assays described in Example 5 are presented. Binding as compared to the standard peptide is shown in the far right column. The column labeled "Pos." indicates the position in the antigenic protein at which the sequence occurs.

[00160] Tables 15 [3] and 16 [4] provide the results of these searches. Binding affinities are expressed as percentage of binding compared to standard peptide in the assays as described in the parent applications are presented.

[00161] Although the present invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims.

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0841	ILSPFLPLL	9	HBV	adr	ENV	371	2.9
1.0240	TLQDIVLHL	9	HPV	18	E7	7	0.76
1.0838	WLSLLVPFV	9	HBV	adr	ENV	335	0.72
1.0851	FLLSLGIHL	9	HBV	adr	POL	1147	0.52
1.0306	QLFEDNYAL	9	c-ERB2			106	0.46
1.0814	LMVTVYYGV	9	HIV		ENV	2182	0.44
1.0878	MMWFWGPSL	9	HBV	adw	ENV	360	0.41
1.0839	MMWYWGPSL	9	HBV	adr	ENV	360	0.41
1.0384	FLTKQYLNL	9	HBV	adw	POL	1279	0.29
1.0321	ILHNGAYSL	9	c-ERB2			435	0.21
1.0834	LLLCLIFLL	9	HBV	adr	ENV	250	0.19
1.0167	GLYSSTVPV	9	HBV	adr	POL	635	0.15
1.0849	HLYSHPIIL	9	HBV	adr	POL	1076	0.13
1.0275	RMPEAAPPV	9	p53			65	0.12
1.0854	LLMGTLGIV	9	HPV	16	E7	82	0.11
1.0880	ILSPTMPLL	9	HBV	adw	ENV	371	0.11
1.0127	YLVAYQATV	9	HCV		LORF	1585	0.11
1.0151	VLLDYQGML	9	HBV	adr	ENV	259	0.11
1.0018	VLAEAMSQV	9	HIV		GAG	367	0.11
1.0330	RLLQETELV	9	c-ERB2			689	0.091
1.0209	SLYAVSPSV	9	HBV	adr	POL	1388	0.078
1.0816	DLMGYIPLV	9	HCV		CORE	132	0.055
1.0835	LLCLIFLLV	9	HBV	adr	ENV	251	0.049
1.0852	FLCQQYLHL	9	HBV	adr	POL	1250	0.048
1.0882	NLYVSLMLL	9	HBV	adw	POL	1088	0.046
1.0637	GMLPVCPLL	9	HBV	adr	ENV	265	0.046
1.0819	ILPCSFTTL	9	HCV		NS1/ENV2	676	0.045
1.0109	ALSTGLIHL	9	HCV		NS1/ENV2	686	0.042
1.0833	ILLLCLIFL	9	HBV	adr	ENV	249	0.035
1.0301	HLYQGCQVV	9	c-ERB2			48	0.034
1.0337	CLTSTVQLV	9	c-ERB2			789	0.034
1.0842	PLLPIFFCL	9	HBV	adr	ENV	377	0.031
1.0861	ALCRWGLLL	9	c-ERB2			5	0.031
1.0309	VLQRNTQL	9	c-ERB2			153	0.029
1.0828	VLQAGFFLL	9	HBV	adr	ENV	177	0.024
1.0844	LLWFHISCL	9	HBV	adr	CORE	490	0.024
1.0135	ILAGYGAGV	9	HCV		LORF	1851	0.024
1.0870	QLMPYGCLL	9	c-ERB2			799	0.023
1.0075	LLWKGEGAV	9	HIV		POL	1496	0.023
1.0873	FLGGTPVCL	9	HBV	adw	ENV	204	0.021
1.0323	ALIHHNTHL	9	c-ERB2			466	0.021
1.0859	VLVHPQWVL	9	PSA			49	0.020
1.0267	KLQCVDLHV	9	PSA			166	0.019
1.0820	VLPCSFTTL	9	HCV		NS1/ENV2	676	0.017
1.0111	HLHQNIVDV	9	HCV		NS1/ENV2	693	0.016
1.0103	SMVGNWAKV	9	HCV		ENV1	364	0.016
1.0283	LLGRNSFEV	9	p53			264	0.014
1.0207	GLYRPLLSL	9	HBV	adr	POL	1370	0.014
1.0389	GLYRPLLRL	9	HBV	adw	POL	1399	0.014
1.0185	NLSWLSLDV	9	HBV	adr	POL	996	0.013
1.0113	FLLLADARV	9	HCV		NS1/ENV2	725	0.013
1.0119	YLVTRHADV	9	HCV		LORF	1131	0.011

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0846	CLTHIVNLL	9	HBV	adr	POL	912	0.010
1.0156	ELMNLATWV	9	HBV	adr	CORE	454	0.010
1.0236	KLPDLCTEL	9	HPV	18	E6	13	0.010
1.0056	ALQDSGLEV	9	HIV		POL	1180	0.0083
1.0375	LLSSDLSWL	9	HBV	adw	POL	1021	0.0081
1.0094	ALAHGVRVL	9	HCV		CORE	150	0.0072
1.0129	TLHGITPLL	9	HCV		LORF	1617	0.0070
1.0041	KLLRGTKAL	9	HIV		POL	976	0.0069
1.0131	CMSADLEVV	9	HCV		LORF	1648	0.0067
1.0872	GLLGFLLVL	9	HBV	adw	ENV	170	0.0066
1.0228	TLHEYMLDL	9	HPV	16	E7	7	0.0059
1.0274	KLLPENNVL	9	p53			24	0.0058
1.0043	ILKEPVHGV	9	HIV		POL	1004	0.0055
1.0206	RLGLYRTLL	9	HBV	adr	POL	1368	0.0050
1.0188	GLPRYVARL	9	HBV	adr	POL	1027	0.0050
1.0202	KLIGTDNSV	9	HBV	adr	POL	1317	0.0050
1.0818	FLLALLSCL	9	HCV		CORE	177	0.0046
1.0184	LLSSNLSWL	9	HBV	adr	POL	992	0.0046
1.0102	QLLRIPQAV	9	HCV		ENV1	337	0.0039
1.0114	GLRDLAVAV	9	HCV		LORF	963	0.0034
1.0005	TLNAWVKVI	9	HIV		GAG	156	0.0032
1.0183	NLQSLTNLL	9	HBV	adr	POL	985	0.0025
1.0359	QLGRKPTPL	9	HBV	adw	ENV	89	0.0025
1.0150	SLDSWWTSL	9	HBV	adr	ENV	194	0.0023
1.0362	ILSKTGDPV	9	HBV	adw	ENV	153	0.0021
1.0866	ILLVVVLGV	9	c-ERB2			661	0.0020
1.0214	LLHKRTLGL	9	HBV	adr	"X"	1510	0.0019
1.0216	CLFKDWEEL	9	HBV	adr	"X"	1533	0.0019
1.0862	GLGISWLGL	9	c-ERB2			447	0.0018
1.0187	HLLVGSSGL	9	HBV	adr	POL	1020	0.0018
1.0318	TLEEITGYL	9	c-ERB2			402	0.0018
1.0328	FLTSIISAV	9	c-ERB2			650	0.0015
1.0822	LLGCIITSL	9	HCV		LORF	1089	0.0015
1.0277	ALNKMFCQL	9	p53			129	0.0013
1.0066	HLEGKIILV	9	HIV		POL	1322	0.0010
1.0308	QLRSLTEIL	9	c-ERB2			141	0.0008
1.0115	DLAVAVEPV	9	HCV		LORF	966	0.0008
1.0391	VLHKRTLGL	9	HBV	adw	"X"	1539	0.0007
1.0876	FLCILLLCL	9	HBV	adw	ENV	246	0.0007
1.0148	LLDPRVRGL	9	HBV	adr	ENV	120	0.0006
1.0221	KLPQLCTEL	9	HPV	16	E6	18	0.0006
1.0065	HLEGKVILV	9	HIV		POL	1322	0.0006
1.0017	EMMTACQGV	9	HIV		GAG	350	0.0006
1.0055	HLALQDSGL	9	HIV		POL	1178	0.0005
1.0868	VLGVVFGIL	9	c-ERB2			666	0.0005
1.0004	TLNAWVKVV	9	HIV		GAG	156	0.0005
1.0381	HLESLYAAV	9	HBV	adw	POL	1165	0.0005
1.0128	CLIRLKPTL	9	HCV		LORF	1610	0.0004
1.0255	CLGLSYDGL	9	MAGE	1/3		174	0.0004
1.0212	HLSLRGLPV	9	HBV	adr	"X"	1470	0.0004
1.0247	ILESLFRAV	9	MAGE	1		93	0.0004
1.0092	TLTCGFADL	9	HCV		CORE	125	0.0003

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0108	TLPALSTGL	9	HCV		NS1/ENV2	683	0.0003
1.0294	ALAIPQCRL	9	EBNA1			525	0.0003
1.0101	DLCCSVFLV	9	HCV		ENV1	280	0.0003
1.0231	RLCVQSTHV	9	HPV	16	E7	66	0.0003
1.0162	LLDDEAGPL	9	HBV	adr	POL	587	0.0002
1.0829	CLRRFIIFL	9	HBV	adr	ENV	239	0.0002
1.0126	CLPVCQDHL	9	HCV		LORF	1547	0.0001
1.0163	PLEEELPRL	9	HBV	adr	POL	594	0.0001
1.0130	FLLYRLGAV	9	HCV		LORF	1623	0.0001
1.0042	ELAENREIL	9	HIV		POL	997	0
1.0054	ELQAIHLAL	9	HIV		POL	1173	0
1.0089	LIPRRGPRL	9	HCV		CORE	36	0
1.0091	NLGKVIDTL	9	HCV		CORE	118	0
1.0093	PLGGAARAL	9	HCV		CORE	143	0
1.0154	DLLDTASAL	9	HBV	adr	CORE	419	0
1.0178	QLKQSRLGL	9	HBV	adr	POL	791	0
1.0179	GLQPQQGSL	9	HBV	adr	POL	798	0
1.0286	PLDGEYFTL	9	p53			322	0
1.0296	VLKDAIKDL	9	EBNA1			574	0
1.0310	QLCYQDTIL	9	c-ERB2			160	0
1.0007	DLNTMLNTV	9	HIV		GAG	188	0
1.0037	ELHPDKWTV	9	HIV		POL	928	0
1.0070	ELKKIIGQV	9	HIV		POL	1412	0
1.0157	ELVVSYVNV	9	HBV	adr	CORE	473	0
1.0160	CLTPGRETV	9	HBV	adr	CORE	497	0
1.0164	DLNLGNLNV	9	HBV	adr	POL	614	0
1.0867	LLVVVLGVV	9	c-ERB2			662	0
1.0159	NMGLKIRQL	9	HBV	adr	CORE	482	0
1.0322	SLRELGSGL	9	c-ERB2			457	<0.0002
1.0350	DLLEKGERL	9	c-ERB2			933	<0.0002
1.0352	DLVDAEEYL	9	c-ERB2			1016	<0.0002
1.0366	FLEEELPIIL	9	HBV	adw	POL	623	<0.0002
1.0372	DLQHGRLVL	9	HBV	adw	POL	781	<0.0002
1.0390	FLPGPLGAL	9	HBV	adw	"X"	1476	<0.0002
1.0811	LLTQIGCTL	9	HIV		POL	685	<0.0002
1.0812	PLVKLWYQL	9	HIV		POL	1116	<0.0002
1.0832	FLFILLLCL	9	HBV	adr	ENV	246	<0.0002
1.0847	NLYVSLLLL	9	HBV	adr	POL	1059	<0.0002
1.0316	PLQPEQLQV	9	c-ERB2			391	<0.0002
1.0342	DLAARNVLV	9	c-ERB2			845	<0.0002
1.0343	VLVKSPNHV	9	c-ERB2			851	<0.0002
1.0356	TLSPGKNGV	9	c-ERB2			1172	<0.0002
1.0376	DLSWLSLDV	9	HBV	adw	POL	1025	<0.0002
1.0363	NMENTASGL	9	HBV	adw	ENV	163	<0.0002
1.0195	TLPQEHIVL	9	HBV	adr	POL	1179	<0.0003
1.0196	KLKQCFRKL	9	HBV	adr	POL	1188	<0.0003
1.0201	PLPIHTAEL	9	HBV	adr	POL	1296	<0.0003
1.0210	QLDPARDVL	9	HBV	adr	"X"	1426	<0.0003
1.0220	VLGGCRHKL	9	HBV	adr	"X"	1561	<0.0003
1.0229	DLQPETTDL	9	HPV	16	E7	14	<0.0003
1.0245	ALEAQQEAL	9	MAGE	1		15	<0.0003
1.0266	DLPTQEPAL	9	PSA			136	<0.0003

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0279	HLIRVEGNL	9	p53			193	<0.0003
1.0262	TLEDSSGNL	9	p53			256	<0.0003
1.0238	ELRHYSDSV	9	HPV	18	E6	77	<0.0003
1.0268	DLFIVISNDV	9	PSA			171	<0.0003
1.0836	CLIFLLVLL	9	HBV	adr	ENV	253	<0.0006

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0890	LLFNILGGWV	10	HCV		LORF	1807	3.5
1.0930	LLVPFVQWFV	10	HBV	adw	ENV	338	1.6
1.0884	LLALLSCLTV	10	HCV		CORE	178	0.61
1.0895	ILLLCLIFLL	10	HBV	adr	ENV	249	0.30
1.0518	GLSPTVWLSV	10	HBV	adr	ENV	348	0.28
1.0902	SLYNILSPFL	10	HBV	adr	ENV	367	0.23
1.0892	LLVLQAGFFL	10	HBV	adr	ENV	175	0.21
1.0686	FLQTHIFAEV	10	EBNA1			565	0.17
1.0628	QLFLNTLSFV	10	HPV	18	E7	88	0.11
1.0904	LLPIFFCLWV	10	HBV	adr	ENV	378	0.10
1.0897	LLLCLIFLLV	10	HBV	adr	ENV	250	0.099
1.0516	LLDYQGMLPV	10	HBV	adr	ENV	260	0.085
1.0901	WMMWYWGPSL	10	HBV	adr	ENV	359	0.084
1.0533	GLYSSTVPVL	10	HBV	adr	POL	635	0.080
1.0469	YLLPRRGPRL	10	HCV		CORE	35	0.073
1.0888	GLLGCIITSL	10	HCV		LORF	1038	0.061
1.0907	ILCWGELMNL	10	HBV	adr	CORE	449	0.052
1.0927	LLGICLTSTV	10	c-ERB2			785	0.049
1.0452	LLWKGEGAVV	10	HIV		POL	1496	0.036
1.0885	LLALLSCLTI	10	HCV		CORE	178	0.034
1.0620	KLTNTGLYNL	10	HPV	18	E6	92	0.032
1.0502	RLIVFPDLGV	10	HCV		LORF	2578	0.032
1.0659	FLTPKKLQCV	10	PSA			161	0.031
1.0932	WMMWFWGPSL	10	HBV	adw	ENV	359	0.029
1.0772	SLNFLGGTPV	10	HBV	adw	ENV	201	0.027
1.0609	SLQDIEITCV	10	HPV	18	E6	24	0.025
1.0525	ILSTLPETTV	10	HBV	adr	CORE	529	0.022
1.0508	RLHGLSAFSL	10	HCV		LORF	2885	0.020
1.0493	ILGGWVAAQL	10	HCV		LORF	1811	0.018
1.0738	VMACVGSPYV	10	c-ERB2			773	0.018
1.0460	QLMVTVYYGV	10	HIV		ENV	2181	0.017
1.0573	ILRGTSFVYV	10	HBV	adr	POL	1345	0.016
1.0703	SLTEILKGGV	10	c-ERB2			144	0.015
1.0912	LLGCAANWIL	10	HBV	adr	POL	1337	0.014
1.0798	ALPPASPSAV	10	HBV	adw	"X"	1483	0.013
1.0908	QLLWFHISCL	10	HBV	adr	CORE	489	0.013
1.0677	NLLGRNSFEV	10	p53			263	0.013
1.0889	VLAALAAYCL	10	HCV		LORF	1666	0.011
1.0528	LLLDDEAGPL	10	HBV	adr	POL	586	0.011
1.0500	LMAKNSVFCV	10	HCV		LORF	2558	0.0088
1.0492	VLVGGVLAAL	10	HCV		LORF	1661	0.0084
1.0898	LLCLIFLLVL	10	HBV	adr	ENV	251	0.0075
1.0458	KLMVTVYYGV	10	HIV		ENV	2181	0.0069
1.0459	NLMVTVYYGV	10	HIV		ENV	2181	0.0067
1.0530	CLSPTVWLSA	10	HBV	adw	ENV	348	0.0067
1.0759	SLPTHDPSPL	10	c-ERB2			1100	0.0059
1.0419	VLPEKDSWTV	10	HIV		POL	940	0.0056
1.0666	FLHSGTAKSV	10	p53			113	0.0050
1.0473	GLIHLHQNIV	10	HCV		NS1/ENV2	690	0.0047
1.0792	SLYAAVTNFL	10	HBV	adw	POL	1168	0.0046
1.0780	IMLARFYPNV	10	HBV	adw	POL	713	0.0043
1.0507	YLTRDPTTPL	10	HCV		LORF	2803	0.0042

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0914	GLYNLLIRCL	10	HIV	18	E6	97	0.0036
1.0649	YLEYGRCRTV	10	MAGE	1		248	0.0034
1.0561	SLFTSITNFL	10	HBV	adr	POL	1139	0.0034
1.0788	NLLSSDLSWL	10	HBV	adw	POL	1020	0.0032
1.0753	RMARDPQRFV	10	c-ERB2			978	0.0020
1.0568	RMRGTFVVTL	10	HBV	adr	POL	1288	0.0020
1.0642	SLQLVFGIDV	10	MAGE	1		150	0.0020
1.0582	KLLHKRTLGL	10	HBV	adr	'X'	1509	0.0019
1.0713	GLGMEHLREV	10	c-ERB2			344	0.0017
1.0742	GMSYLEDVRL	10	c-ERB2			832	0.0017
1.0549	NLLSSNLSWL	10	HBV	adr	POL	991	0.0016
1.0465	QLTVWGIKQL	10	HIV		ENV	2760	0.0015
1.0524	VLEYLVSFGV	10	HBV	adr	CORE	505	0.0015
1.0483	VLNITSVAATL	10	HCV		LORF	1253	0.0015
1.0548	SLTNLLSSNL	10	HBV	adr	POL	988	0.0014
1.0512	ALLDPRVRGL	10	HBV	adr	ENV	119	0.0011
1.0676	TLEDSSGNLL	10	p53			256	0.0011
1.0719	TLQGLGISWL	10	c-ERB2			444	0.0011
1.0627	DLRAPQQLFL	10	HPV	18	E7	82	0.0010
1.0725	VLQGLPREYV	10	c-ERB2			546	0.0009
1.0918	DLPPWFPPMV	10	EBNA1			605	0.0009
1.0499	DLSDGSWSTV	10	HCV		LORF	2399	0.0008
1.0559	CLAPSYMDDV	10	HBV	adr	POL	1118	0.0008
1.0632	PLVLGTLEEV	10	MAGE	1		37	0.0008
1.0520	NLATWVGSNL	10	HBV	adr	CORE	457	0.0008
1.0400	NLLTQIGCTL	10	HIV		POL	684	0.0007
1.0488	GLTHIDAHFL	10	HCV		LORF	1564	0.0007
1.0733	VLGSGAFGTV	10	c-ERB2			725	0.0007
1.0434	QLIKKEKVYL	10	HIV		POL	1219	0.0006
1.0451	KLLWKGEGAV	10	HIV		POL	1495	0.0006
1.0470	SMVGNWAKVL	10	HCV		ENV1	364	0.0006
1.0570	KLIGTDNSVV	10	HBV	adr	POL	1317	0.0006
1.0924	TLLVVVLGVV	10	c-ERB2			661	0.0006
1.0397	LLDTGADDTV	10	HIV		POL	619	0.0005
1.0446	HLKTAVQMAV	10	HIV		POL	1426	0.0005
1.0601	DLLMGTLGIV	10	HPV	16	E7	81	0.0005
1.0443	LLKLAGRWPV	10	HIV		POL	1356	0.0004
1.0461	DLMVTVYYGV	10	HIV		ENV	2181	0.0004
1.0619	TLEKLTNTGL	10	HPV	18	E6	89	0.0004
1.0787	SLTNLLSSDL	10	HBV	adw	POL	1017	0.0004
1.0521	NLEDPASREL	10	HBV	adr	CORE	465	0.0003
1.0583	GLSAMSTTDL	10	HBV	adr	'X'	1517	0.0003
1.0652	VLVASRGRAV	10	PSA			36	0.0003
1.0716	DLSVFQNLQV	10	c-ERB2			421	0.0003
1.0723	QLFRNPHQAL	10	c-ERB2			484	0.0003
1.0727	PLTSIISAVV	10	c-ERB2			650	0.0003
1.0479	YLKGSSGGPL	10	HCV		LORF	1160	0.0002
1.0497	QLIPEPEPDV	10	HCV		LORF	2159	0.0002
1.0523	CLTFGRETVL	10	HBV	adr	CORE	497	0.0002
1.0603	TLEDLLMGTL	10	HPV	16	E7	78	0.0002
1.0631	SLHCKPEEAL	10	MAGE	1		7	0.0002
1.0680	EMPRELNEAL	10	p53			339	0.0002

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0689	VLKDAIKDLV	10	EBNA1			574	0.0002
1.0757	DLVDAEEYLV	10	c-ERB2			1016	0.0002
1.0796	RMRGTFVSFL	10	HBV	adw	POL	1317	0.0002
1.0669	QILAKTCPVQL	10	p53			136	0.0001
1.0717	NLQVIRGRIL	10	c-ERB2			427	0.0001
1.0721	WLGLRSLREL	10	c-ERB2			452	0.0001
1.0522	NMGLKIRQLL	10	HBV	adr	CORE	482	0
1.0527	PLSYQHFRKL	10	HBV	adr	POL	576	0
1.0529	ELFRLADEGL	10	HBV	adr	POL	598	0
1.0531	GLNRRVAEDL	10	HBV	adr	POL	606	0
1.0536	PLTVNEKRRL	10	HBV	adr	POL	672	0
1.0539	IMPARFYFNL	10	HBV	adr	POL	684	0
1.0550	PLHPAAMPHL	10	HBV	adr	POL	1012	0
1.0552	DLHDSCSRNL	10	HBV	adr	POL	1051	0
1.0555	LLYKTFGRKL	10	HBV	adr	POL	1066	0
1.0557	PMGVGLSPFL	10	HBV	adr	POL	1090	0
1.0560	VLGAKSVQHL	10	HBV	adr	POL	1128	0
1.0569	PLPIHTAELL	10	HBV	adr	POL	1296	0
1.0579	PLPSLAPSAV	10	HBV	adr	'X'	1454	0
1.0583	DLEAYFKDCL	10	HBV	adr	'X'	1525	0
1.0587	ELGEEIRLKV	10	HBV	adr	'X'	1540	0
1.0589	VLGGCRHKLV	10	HBV	adr	'X'	1551	0
1.0597	TLEQQYNKPL	10	HPV	16	E6	94	0
1.0608	DLCTELNTSL	10	HPV	18	E6	16	0
1.0616	RLQRRRETQV	10	HPV	18	E6	49	0
1.0621	HLEPQNEIPV	10	HPV	18	E7	14	0
1.0639	LLKYRAREPV	10	MAGE	1/3		114	0
1.0643	CLGLSYDGLL	10	MAGE	1/3		174	0
1.0657	DMSLLKNRFL	10	PSA			98	0
1.0658	LLRLSEPAEL	10	PSA			119	0
1.0663	PLSQETFSDL	10	p53			13	0
1.0664	PLPSQAMDDL	10	p53			34	0
1.0690	ELAALCRWGL	10	c-ERB2			2	0
1.0692	RLPASPETFL	10	c-ERB2			34	0
1.0699	RLRIVRGTQL	10	c-ERB2			98	0
1.0701	GLRELQLRSL	10	c-ERB2			136	0
1.0730	QMRILKETEL	10	c-ERB2			711	0
1.0732	ILKETELRKV	10	c-ERB2			714	0
1.0754	PLDSTFYRSL	10	c-ERB2			999	0
1.0755	LLEDDDMGDL	10	c-ERB2			1008	0
1.0758	DLGMGAAKGL	10	c-ERB2			1089	0
1.0761	PLPSETDGYV	10	c-ERB2			1119	0
1.0763	TLSPGKNGVV	10	c-ERB2			1172	0
1.0765	TLQDPRVRAL	10	HBV	adw	ENV	119	0
1.0768	NMENIASGLL	10	HBV	adw	ENV	163	0
1.0775	ELFILADEGL	10	HBV	adw	POL	627	0
1.0776	GLNRPVAEDL	10	HBV	adw	POL	635	0
1.0777	PLTVNENRRL	10	HBV	adw	POL	701	0
1.0790	LLYKTYGRKL	10	HBV	adw	POL	1095	0
1.0801	GLSAMSPTDL	10	HBV	adw	'X'	1546	0
1.0802	DLEAYFKDCV	10	HBV	adw	'X'	1554	0
1.0803	TLQDPRVRGL	10	HBV	ayw	ENV	119	0

Table 14

Peptide	Sequence	AA	Virus	Strain	Molecule	Pos.	A2.1
1.0804	NMENITSGFL	10	HBV	ayw	ENV	163	0
1.0891	DLVNLLPAIL	10	HCV		LORF	1578	0
1.0404	PLTEEKIKAL	10	HIV		POL	720	<0.0002
1.0409	QLGIPHPAGL	10	HIV		POL	786	<0.0002
1.0411	GLKKKKSVTV	10	HIV		POL	794	<0.0002
1.0450	PIWKGPAKLL	10	HIV		POL	1488	<0.0002
1.0476	DLAVAVEPVV	10	HCV		LORF	966	<0.0002
1.0478	SLTGRDKNQV	10	HCV		LORF	1046	<0.0002
1.0490	DLEVVTSTWV	10	HCV		LORF	1652	<0.0002
1.0494	GLGKVLIDIL	10	HCV		LORF	1843	<0.0002
1.0505	VLTTSCGNTL	10	HCV		LORF	2704	<0.0002
1.0506	ELITSCSSNV	10	HCV		LORF	2781	<0.0002
1.0510	CLRKLGVPPL	10	HCV		LORF	2908	<0.0002
1.0511	PLGFFPDHQL	10	HBV	adr	ENV	10	<0.0002
1.0514	NMENTTSGFL	10	HBV	adr	ENV	163	<0.0002

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
ALEAQQEAL	9	1		15	2.1		<0.0003			
ILESLFRAV	9	1		93	2.1		0.0004			
VITKKVADL	9	1		101	2.1		<0.0003			
CLGLSYDGL	9	1/3		174	2.1		0.0004			
QIMPKTGFL	9	1		187	2.1		0.0007			
SLHCKPEEAL	10	1		7	2.1		0.0002			
PLVLGTLEEV	10	1		37	2.1		0.0008			
CILESLFRAV	10	1		92	2.1		0.0003			
AVITKKVADL	10	1		100	2.1		0			
VITKKVADLV	10	1		101	2.1		0			
LLKYRAREPV	10	1/3		114	2.1		0			
EIFGKASESL	10	1		142	2.1		0			
CLGLSYDGLL	10	1/3		174	2.1		0			
AISRKMVEL	9	2		101	2.1		0.0003			
KMVELVHFL	9	2		105	2.1		0.16			
MVELVHFLL	9	2		106	2.1		0.0031			
DLQQSLRVL	9	2		143	2.1		0			
SLRVLAAGL	9	2		147	2.1		0.0001			
ALSRKVAEL	9	3		101	2.1		0.0050			
HLYIFATCL	9	3		167	2.1		0.0003			
YIFATCLGL	9	3		169	2.1		0.018			
QIMPKAGLL	9	3		187	2.1		0			

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
AISRKMVELV	10	2		101	2.1		0			
MVELVHFLLL	10	2		106	2.1		0.0017			
KLPGLLSRDL	10	2		135	2.1		0			
LLSRDLQQSL	10	2		139	2.1		0.0007			
SLPTTMNYPL	10	3		63	2.1		0.0035			
DLESEFQAAL	10	3		93	2.1		0.0001			
ALSRKVAELV	10	3		101	2.1		0.0001			
KVAELVHFLL	10	3		105	2.1		0.012			
VIFSKASSSL	10	3		142	2.1		0			
SLQLVFGIEL	10	3		150	2.1		0.0049			
LMEVDPIGHL	10	3		159	2.1		0.0005			
FLIIVLVMI	9	1		194	2.1		0.0005			
GLLGDNQIM	9	1		181	2.1		0.0051			
SLHCKPEEA	9	1		7	2.1		0.013	<0.0002	0	
ALGLVCVQA	9	1		22	2.1		0.015	<0.0002	<0.0002	
CKPEEALEA	9	1		10	Random		<0.0002			
QQEALGLVC	9	1		19	Random		<0.0002			
VQAATSSSS	9	1		28	Random		<0.0002			
PLVLGTLEE	9	1		37	Random		<0.0002			
VPTAGSTDP	9	1		46	Random		<0.0002			
PQSPQGASA	9	1		55	Random		<0.0002			
FPTTINFTR	9	1		64	Random		<0.0002			

Table 15

Sequence	AA	Mag Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
QRQPSEGSS	9	1		73	Random		<0.0002			
SREEEGPST	9	1		82	Random		<0.0002			
AVITKKVAD	9	1		100	Random		<0.0002			
EMLESVIKN	9	1		127	Random		<0.0002			0
YKHCFPEIF	9	1		136	Random		<0.0002			
GKASESLQL	9	1		145	Random		<0.0002			
VFGIDVKEA	9	1		154	Random		<0.0002	<0.0002	0	
DPTGHSYVL	9	1		163	Random		<0.0002			
VTCLGLSYD	9	1		172	Random		<0.0002			
PKTGFLIIV	9	1		190	Random		<0.0002			
LVMIAMEGG	9	1		199	Random		<0.0002			
HAPEEEIWE	9	1		208	Random		<0.0002			
ELSVMEVYD	9	1		217	Random		<0.0002			
GREHSAYGE	9	1		226	Random		<0.0002			
PRKLLTQDL	9	1		235	Random		0.0002			
VQEKYLEYG	9	1		244	Random		<0.0002			
RCRTVIPHA	9	1		253	Random		<0.0002			
MSSCGVQGP	9	1		262	Random		<0.0002			
ILESLFRAVI	10	1		93	2.1		0.0002			
FLIIVLVMIA	10	1		194	2.1		0.0003	0.0093	0.0030	
LVFGIDVKEA	10	1		153	2.1		0.0002	<0.0002	0	
EVYDGREHSA	10	1		222	2.1		0	<0.0002	0	

Table 15

Sequence	AA	Mag Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
GVQGPSLKPA	10	1		266	2.1		0.0001			
QLVFGIDV	8	1		152	2.1		0			
KLLTQDLV	8	1		237	2.1		0.0004			
GLLGDNQI	8	1		181	2.1		0			
DLVGFLLL	8	1		108	2.1		0			
GLSYDGLL	8	1		176	2.1		0.0001			
DLVQEKYL	8	1		242	2.1		0			
LLGDNQIM	8	1		182	2.1		0			
FLIIVLVM	8	1		194	2.1		0			
ALEAQQEA	8	1		15	2.1		0			
TLEEVPTA	8	1		42	2.1		0			
IMPKTGFL	8	1		188	2.1		0.0001			
PVTKAEML	8	1		122	2.1		0			
IVLVMIAM	8	1		197	2.1		0.0001			
AVITKKVA	8	1		100	2.1		0			
EIWEELSV	8	1		213	2.1		0			
LIIVLVMI	8	1		195	2.1		0.0001			
IIVLVMIA	8	1		196	2.1		0.0002			
SLFRAVITKKV	11	1		96	2.1		0.0001			
LLLKYRAREPV	11	1		113	2.1		0.0001			
YLEYGRCRTVI	11	1		248	2.1		0.0006			
ALEAQQEALGL	11	1		15	2.1		0.0001			

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
FLIIVLVMIAM	11	1		194	2.1		0.0041			
VLGTLEEVPTA	11	1		39	2.1		0.0002			
QLVFGIDVKEA	11	1		152	2.1		0.0001			
AVITKKVADLV	11	1		100	2.1		0			
PVTKAEMLESV	11	1		122	2.1		0			
KVADLVGFLLL	11	1		105	2.1		0.020			
GVQGPSLKPAM	11	1		266	2.1		0			
LVGFLLLKYRA	11	1		109	2.1		0.0004			
LVMIAMEGGHA	11	1		199	2.1		0.0005			
CILESLFRAVI	11	1		92	2.1		0.0030			
EALEAQQEA	9	1		14	2.1		0	<0.0002	0	
EAQQEALGL	9	1		17	2.1		0			<0.0002
AATSSSSPL	9	1		30	2.1		0			<0.0002
ATSSSSPLV	9	1		31	2.1		0.0007			
GTLEEVPTA	9	1		41	2.1		0.013	<0.0002	0	
GASAFPTTI	9	1		60	2.1		0			<0.0002
STSCILESL	9	1		89	2.1		0.0002			
RAVITKKVA	9	1		99	2.1		0	<0.0002	0	
ITKKVADLV	9	1		102	2.1		0			
RAREPVTKA	9	1		118	2.1		0			
KAEMLESVI	9	1		125	2.1		0			<0.0002
KASESLQLV	9	1		146	2.1		0.0009			

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
PTGHSYVLV	9	1		164	2.1		0			
KTGFLIIVL	9	1		191	2.1		0.0006			
LIIVLVMIA	9	1		195	2.1		0	0.0022	0.0006	
IIVLVMIAM	9	1		196	2.1		0.0007			
MIAMEGGHA	9	1		201	2.1		0.0005	<0.0002	0.0002	
EIWEELSVM	9	1		213	2.1		0			
SAYGEPRKL	9	1		230	2.1		0.0002			<0.0002
YLEYGRCRT	9	1		248	2.1		0			
EALGLVCVQA	10	1		21	2.1		0.0005	<0.0002	0	
QAATSSSSPL	10	1		29	2.1		0			<0.0002
VTKAEMLESV	10	1		123	2.1		0			
EADPTGHSYV	10	1		161	2.1		0			
VLGTLEEVPT	10	1		39	2.1		0.0004			
SAFPTTINFT	10	1		62	2.1		0			
GIDVKEADPT	10	1		156	2.1		0			
PTGHSYVLVT	10	1		164	2.1		0			
FLWGPRALA	9	1	new	265	2.1		0.042	0.0017	0	
LAETSYVKV	9	1	new	272	2.1		0			
YVKVLEYVI	9	1	new	277	2.1		0.0002			
RVRFFFPSL	9	1	new	290	2.1		0.0001			
LAETSYVKVL	10	1	new	272	2.1		0			<0.0002
VLEYVIKVSA	10	1	new	280	2.1		0.0002	0.0002	0	

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
AALREEEEGV	10	1	new	301	2.1		0			
SMHCKPEEV	9	1	new (a)	7	2.1		0.018			
AMGLVCVQV	9	1	new (a)	22	2.1		0.012			
LMLGTLEEV	9	1	new (a)	38	2.1		0.13			
LQLVFGIDV	9	1	new	151	2.1		0.0004			
GLSYDGLLG	9	1	new	176	2.1		0			
GLSYDGLLV	9	1	new (a)	176	2.1		0.0047			
LLGDNQIMP	9	1	new	182	2.1		0.0001			
LLGDNQIMV	9	1	new (a)	182	2.1		0.043			
WEELSVMEV	9	1	new	215	2.1		0			
WMELSVMEV	9	1	new (a)	215	2.1		0.041			
RKLLTQDLV	9	1	new	236	2.1		0			
YEFLWGPRA	9	1	new	262	2.1		0			
YMFLWGPRV	9	1	new (a)	262	2.1		0.22			
AATSSSSPLV	10	1	new	30	2.1		0			
ATSSSSPLVL	10	1	new	31	2.1		0			
KMADLVGFLV	10	1	new (a)	105	2.1		1.5			
VADLVGFLLL	10	1	new	106	2.1		0.0008			0.0003
SESLQLVFGI	10	1	new	148	2.1		0			
VMVTCLGLSV	10	1	new (a)	170	2.1		0.30			
QIMPKTGFLI	10	1	new	187	2.1		0.0009			
QMMPKTGFLV	10	1	new (a)	187	2.1		0.050			

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
KTGFLIIVLV	10	1	new	191	2.1		0.0012			
LIIVLVMIAM	10	1	new	195	2.1		0.0003			
VMIAMEGGHV	10	1	new (a)	200	2.1		0.053			
SAYGEPRKLL	10	1	new	230	2.1		0			0.0008
ALAETSYVKVL	11	1 N		270	2.1		0.012			
KMVELVHFLLL	11	2		52	2.1		0.67			
ELMEVDPIGHL	11	3		105	2.1		0.026			
HLYIFATCLGL	11	3		114	2.1		0.041			
LLLKYRAREPV	11	3		60	2.1		0.0001			
QLVFGIELMEV	11	3		99	2.1		0.34			
IMPKAGLLIIV	11	3		135	2.1		0.013			
VLVTCLGLSYDGL	13	1 n	E6	170	2.1		0.0017			
KLLTQDLVQEKYL	13	1 n	E6	237	2.1		0.0060			
DLVQEKYLEYRQV	13	1 n	E6	242	2.1		0			
SLFRAVITKKVADLV	15	1 n	POL	96	2.1		0.0004			
DLESEFQAAISRKMV	15	2	POL	40	2.1		0			
MLGSVVGNWQYFFPV	15	3	POL	75	2.1		0.012			
GASSFSTTI	9	2		60	2.1		0			0.0002
DLESEFQAA	9	2,3		93	2.1		0			
QAAISRKMV	9	2		99	2.1		0			
KAEMLESVL	9	2		125	2.1		0			0
KASEYLQLV	9	2		146	2.1		0.011			

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
QLVFGIEVV	9	2		152	2.1		0.0038			
VVPISHLYI	9	2		162	2.1		0.0002			
PISHLYILV	9	2		164	2.1		0.0005			
HLYILVTCL	9	2		167	2.1		0.0034			
YILVTCLGL	9	2		169	2.1		0.0014			
GLLGDNQVM	9	2		181	2.1		0.0038			
QVMPKTGLL	9	2		187	2.1		0			
VMPKTGLLI	9	2		188	2.1		0.0010			0.230
KTGLLIIVL	9	2		191	2.1		0.0002			
GLLIIVLAI	9	2,3		193	2.1		0.0002			
LLIIVLAII	9	2,3		194	2.1		0.0001			
LIIVLAIIA	9	2,3		195	2.1		0.0008			
IIVLAIIAI	9	2		196	2.1		0.0009			
IIAIEGDCA	9	2		201	2.1		0			
GASSLPTTM	9	3		60	2.1		0			0.0010
QAALSRKVA	9	3		99	2.1		0			
VAELVHFLL	9	3		106	2.1		0			0.039
KAEMLGSVV	9	3		125	2.1		0			
KASSSLQLV	9	3		146	2.1		0.0005			
QLVFGIELM	9	3		152	2.1		0.0010			
PIGHLYIFA	9	3		164	2.1		0			
IMPKAGLLI	9	3		188	2.1		0.0064			

Table 15

Sequence	AA	Mag Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
KAGLLIIVL	9	3		191	2.1		0.0002			0
IIAREGDCA	9	3		201	2.1		0			
EALEAQQEAL	10	1	new	14	2.1		0			0
EAQQEALGLV	10	1	new	17	2.1		0			
DLESEFQAAI	10	2		93	2.1		0			
AAISRKMVEL	10	2		100	2.1		0			0
VIFSKASEYL	10	2		142	2.1		0.0014			
YLQLVFGIEV	10	2		150	2.1		0.37			
LVFGIEVVEV	10	2		153	2.1		0.012			
GIEVVEVVPI	10	2		156	2.1		<0.0002			
VVEVVPISHL	10	2		159	2.1		<0.0002			
EVVPISHLYI	10	2		161	2.1		<0.0002			
VVPISHLYIL	10	2		162	2.1		0.0002			
PISHLYILVT	10	2		164	2.1		0.0003			
QVMPKTGLLI	10	2		187	2.1		0.0002			
VMPKTGLLII	10	2		188	2.1		0.0009			0.058
KTGLLIIVLA	10	2		191	2.1		<0.0002			
GLLIIVLAII	10	2,3		193	2.1		0.0005			
LLIIVLAIIA	10	2,3		194	2.1		<0.0002			
LIIVLAIIAI	10	2		195	2.1		0.0013			
AIIAIEGDCA	10	2		200	2.1		0.0023			
AALSRKVAEL	10	3		100	2.1		0.0007			0

Table 15

S quence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
VAELVHFLLL	10	3		106	2.1		0.0009			0.0010
VTKAEMLGSV	10	3		123	2.1		<0.0002			
GIELMEVDPI	10	3		156	2.1		<0.0002			
EVDPIGHLYI	10	3		161	2.1		<0.0002			
PIGHLYIFAT	10	3		164	2.1		0.0003			
QIMPKAGLLI	10	3		187	2.1		0.0006			
IMPKAGLLII	10	3		188	2.1		0.0015			
KAGLLIIVLA	10	3		191	2.1		<0.0002			
AIIAREGDCA	10	3		200	2.1		<0.0002			
FLWGPRALI	9	2		271	A02					
GLEARGEAL	9	3		15	A02					
EARGEALGL	9	3		17	A02					
ALGLVGAQA	9	3		22	A02/A03					
GLVGAQAPA	9	3		24	A02/A03					
LVGAQAPAT	9	3		25	A02					
PATEEQEAA	9	3		31	A02/A03					
EAASSSSTL	9	3		37	A02					
AASSSSTLV	9	3		38	A02					
LVEVTLGEV	9	3		45	A02					
EVTLGEVPA	9	3		47	A02/A03					
VTLGEVPAA	9	3		48	A02/A03					
KIWEELSVL	9	3		220	A02					

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
SILGDPKKL	9	3		237	A02					
ILGDPKKLL	9	3		238	A02					
FLWGPRALV	9	3		271	A02					
RALVETSYV	9	3		276	A02					
LVETSYVKV	9	3		278	A02					
YVKVLHHMV	9	3		283	A02					
KVLHHMVKI	9	3		285	A02					
EARGEALGLV	10	3		17	A02					
EALGLVGAQA	10	3		21	A02/A03					
GLVGAQAPAT	10	3		24	A02					
QAPATEEQEA	10	3		29	A02/A03					
EAASSSSTLV	10	3		37	A02					
TLVEVTLGEV	10	3		44	A02					
EVTLGEVPAA	10	3		47	A02/A03					
EVFEGREDSI	10	3		229	A02					
SILGDPKKLL	10	3		237	A02					
ILGDPKKLLT	10	3		238	A02					
ALVETSYVKV	10	3		277	A02					
LVETSYVKVL	10	3		278	A02					
MVKISGGPHI	10	3		290	A02					
LVLGTLEEV	9	1		38	2.1	<0.0006	0.032	0	0	0.0003
KVADLVGFLL	10	1		105		0.0005	0.041	0.0039	0.0030	0.0070

Table 15

Sequence	AA	Mag Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
LVFGIELMEV	10	3		153	2.1		0.17			
ILLWQPIPV	9	3				<0.0007	1.4	0.0048	0.0048	0
EVDPIGHLY	9	3				3.7			0.0022	
KMVELVHFL	9	2				<0.0007	0.13	0.0007	0	0.0043
KMVELVHFLL	10	2		105		<0.0008	0.071	0.0004	0.0001	0.0008
LVFGIELMEV	10	3				0.0030	0.065	0.0007	0	0
KVAELVHFL	9	3		105	2.1	0	0.073	0.011	0.0047	0.0005
CILESLFRA	9	1		92	2.1	0.0001	0.073	0	0.0002	0
VMIAMEGGHA	10	1		200	2.1	<0.00008	0.0023	0	0	0
MLESVIKNYK	10	1				0	0	0.034	0.0045	0
ETSYVKVLEY	10	1				0.075	0	0.0009	0.0004	0
KVLEYVIKV	9	1	new	279	2.1	<0.0005	0.095	0.022	0.015	0
FLWGPRALA	9	1				<0.0006	0.027	0.0015	0	0
ALREEEEGV	9	1		302	2.1	<0.0006	0.0056	0	0	0
ALAETSYVKV	10	1		271		<0.0007	0.017	0.0011	0.0029	0
YVIKVSARV	9	1		283	2.1	0.0005	0.018	0	0	0
RALAETSYV	9	1		270	2.1	<0.0006	0.014	0.0003	0.0005	0
ALAETSYVK	9	1				<0.0006	0.0002	0.17	0.39	0
VLGTLEEV	8	1		39	2.1	<0.0007	0.0088	0	0	0
SLQLVFGI	8	1		150	2.1	<0.0007	0.0094	0	0.0001	0
ILESLFRA	8	1		93	2.1	<0.0004	0.0017	0.0003	0	0.0001
FLLLKYRA	8	1		112	2.1	0.0036	0.0007	0.0003	0.0001	0

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
GLVCVQAA	8	1		24	2.1	0.0016	0.0008	0.0008	0	0
VLVTCLGL	8	1		170	2.1	<0.0007	0.0010	0.0001	0	0
KVADLVGFL	9	1		105	2.1	<0.0008	0.0091	0.0013	0.0005	0
YVLVTCLGL	9	1		169	2.1					
IMPKTGFLI	9	1		188	2.1	<0.0008	0.0035	0	0	3.2
GLLGDNQIM	9	1			A2.1	<0.0008	0.0054	0	0	0.0002
GLVCVQAAT	9	1		24	2.1	0.0030	0.0007	0.0026	0	0.0001
VADLVGFLL	9	1		106	2.1	0.032	0.0011	0.0054	0:0008	0.0007
YLEYGRCRTV	10	1		248	2.1	0.0008	0.0097	0.0001	0	0
SLQLVFGIDV	10	1		150	2.1	0.0028	0.0047	0.0013	0.0001	0.0001
IMPKTGFLII	10	1		188	2.1	<0.0008	0.0007	0	0	0.050
ALGLVCVQAA	10	1		22	A2.1	0.0011	0.0002	0.0003	0	0
EIWEELSVMEV	11	1		213	A2.1	0.0007	0.013	0.0001	0.0001	0
FLIIVLVMIAM	11	1			A2.1	0.023	0.0031	0.016	0.0014	0.0011
VIPHAMSSCGV	11	1		257	2.1	<0.0009	1.4	0	0	0
CILESCFRAVI	11	1			A2.1	0.079	0.0017	0.058	0.0005	0.0008
QIMPKTGFLII	11	1		187	2.1	<0.0009	0.0003	0	0	0.0030
GFLLLKYRA	9	1						0.0004	0.0002	
CFPEIFGKA	9	1						0	0	
FFFPSLREA	9	1						0	0	
FFPSLREAA	9	1						0	0	
RSLHCKPEEA	10	1						0.0001	0.0008	

Table 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
EFLWGPRALA	10	1						0	0	
RFFFPSLREA	10	1						0.0004	0	
FFFPSLREAA	10	1						0	0	

F:\PERFSOL\DOCS\OU9202.1

95

Table 16

Sequence	Antigen	Strain	Molecule	Position	Motif	A1	A2	A3	A11	A24	Max.
						Binding	Binding	Binding	Binding	Binding	Binding
ALFLGFLGAA	HIV	MN	gp160	518	A02		0.4950				0.4950
MLQLTVWGI	HIV	MN	gp160	566	A02		0.2450				0.2450
RVIEVLQRA	HIV	MN	gp160	829	A02		0.1963				0.1963
KLTPLCVTL	HIV	MN	gp160	120	A02		0.1600				0.1600
LLIAARIVEL	HIV	MN	gp160	776	A02		0.1550				0.1550
SLLNATDIAV	HIV	MN	gp160	814	A02		0.1050				0.1050
ALFLGFLGA	HIV	MN	gp160	518	A02		0.0945				0.0945
HMLQLTVWGI	HIV	MN	gp160	565	A02		0.0677				0.0677
LLNATDIAV	HIV	MN	gp160	815	A02		0.0607				0.0607
ALLYKLDIV	HIV	MN	gp160	179	A02		0.0362				0.0362
WLWYIKIFI	HIV	MN	gp160	679	A02		0.0355				0.0355
TIIVHLNESV	HIV	MN	gp160	288	A02		0.0350				0.0350
LLQYWSQEL	HIV	MN	gp160	800	A02		0.0265				0.0265
IMIVGGLVGL	HIV	MN	gp160	687	A02		0.0252				0.0252
LLYKLDIVSI	HIV	MN	gp160	180	A02		0.0245				0.0245
FLAIIWVDL	HIV	MN	gp160	753	A02		0.0233				0.0233
TLQCKIKQII	HIV	MN	gp160	415	A02		0.0200				0.0200
GLVGLRIVFA	HIV	MN	gp160	692	A02		0.0195				0.0195
FLGAAGSTM	HIV	MN	gp160	523	A02		0.0190				0.0190
IISLWDQSL	HIV	MN	gp160	107	A02		0.0179				0.0179
TVWGIKQLQA	HIV	MN	gp160	570	A02		0.0150				0.0150
LLGRRGWEV	HIV	MN	gp160	785	A02		0.0142				0.0142
AVLSIVNRV	HIV	MN	gp160	701	A02		0.0132				0.0132

Table 16

Sequence	Antigen	Strain	Molecule	Position	Motif	A1 Binding	A2 Binding	A3 Binding	A11 Binding	A24 Binding	Max. Binding
FIMIVGGLV	HIV	MN	gp160	686	A02		0.0131				0.0131
LLNATDIAVA	HIV	MN	gp160	815	A02		0.0117				0.0117
FLYGALLLA	PLP	Human		80	A02		1.9000				1.9000
SLLTFMIAA	PLP	Human		253	A02		0.5300				0.5300
FMIAATYNFAV	PLP	Human		257	A02		0.4950				0.4950
RMYGVLPWI	PLP	Human		205	A02		0.1650				0.1650
IAATYNFAV	PLP	Human		259	A02		0.0540				0.0540
GLLECCARCLV	PLP	Human		2	A02		0.0515				0.0515
YALTVVWLL	PLP	Human		157	A02		0.0415				0.0415
ALTVVWLLV	PLP	Human		158	A02		0.0390				0.0390
FLYGALLL	PLP	Human		80	A02		0.0345				0.0345
SLCADARMYGV	PLP	Human		199	A02		0.0140				0.0140
LLVFACSAV	PLP	Human		164	A02		0.0107				0.0107

Claims

1. A composition comprising an immunogenic peptide having an HLA-A2.1 binding motif, which immunogenic peptide has 9 residues and the following residues:

a first conserved residue at the second position from the N-terminus selected from the group consisting of I, V, A and T;

a second conserved residue at the C-terminal position selected from the group consisting of V, L, I, A and M.

2. A composition comprising an immunogenic peptide having an HLA-A2.1 binding motif, which immunogenic peptide has 9 residues:

a first conserved residue at the second position from the N-terminus selected from the group consisting of L, M, I, V, A and T;

a second conserved residue at the C-terminal position selected from the group consisting of A and M.

3. The composition of claim 1, wherein the amino acid at position 1 is not an amino acid selected from the group consisting of D, and P.

4. The composition of claim 2, wherein the amino acid at position 1 is not an amino acid selected from the group consisting of D, and P.

5. The composition of claim 0, wherein the amino acid at position 3 from the N-terminus is not an amino acid selected from the group consisting of D, E, R, K and H.

6. The composition of claim 2, wherein the amino acid at position 3 from the N-terminus is not an amino acid selected from the group consisting of D, E, R, K and H.

7. The composition of claim 1, wherein the amino acid at position 6 from the N-terminus is not an amino acid selected from the group consisting of R, K and H.

8. The composition of claim 2, wherein the amino acid at position 6 from the N-terminus is not an amino acid selected from the group consisting of R, K and H.

9. The composition of claim 0, wherein the amino acid at position 7 from the N-terminus is not an amino acid selected from the group consisting of R, K, H, D and E.

10. The composition of claim 2, wherein the amino acid at position 7 from the N-terminus is not an amino acid selected from the group consisting of R, K, H, D and E.

11. A composition comprising an immunogenic peptide having an HLA-A2.1 binding motif, which immunogenic peptide has about 10 residues:

a first conserved residue at the second position from the N-terminus selected from the group consisting of L, M, I, V, A, and T; and

a second conserved residue at the C-terminal position selected from the group consisting of V, I, L, A and M;

wherein the first and second conserved residues are separated by 7 residues.

12. The composition of claim 11, wherein the amino acid at position 1 is not an amino acid selected from the group consisting of D, E and P.

13. The composition of claim 11, wherein the amino acid at position 3 from the N-terminus is not an amino acid selected from the group consisting of D and E.

14. The composition of claim 11, wherein the amino acid at position 4 from the N-terminus is not an amino acid selected from the group consisting of A, K, R and H.

15. The composition of claim 11, wherein the amino acid at position 5 from the N-terminus is not P.

16. The composition of claim 11, wherein the amino acid at position 7 from the N-terminus is not an amino acid selected from the group consisting of R, K and H.

17. The composition of claim 11, wherein the amino acid at position 8 from the N-terminus is not an amino acid selected from the group consisting of D, E, R, K and H.

18. The composition of claim 11, wherein the amino acid at position 9 from the N-terminus is not an amino acid selected from the group consisting of R, K and H.

Abstract of the Disclosure

The present invention provides the means and methods for selecting immunogenic peptides and the immunogenic peptide compositions capable of specifically binding glycoproteins encoded by HLA-A2.1 allele and inducing T cell activation in T cells restricted by the A2.1 allele. The peptides are useful to elicit an immune response against a desired antigen.